

UNIVERSIDAD CARLOS III DE MADRID

Departamento Ingeniería Telemática



PROYECTO FIN DE CARRERA

DISEÑO E IMPLEMENTACIÓN DE UN PORTAL WEB PARA EL ACCESO A LA INFORMACIÓN SOBRE PUBLICADORES EN BITTORRENT

Autora: M^a Almudena González de la Cruz

Tutor: Rubén Cuevas Rumín

Leganés, enero de 2013

Título: DISEÑO E IMPLEMENTACIÓN DE UN PORTAL WEB PARA EL ACCESO
A LA INFORMACIÓN SOBRE PUBLICADORES EN BITTORRENT

Autor: M^a ALMUDENA GONZÁLEZ DE LA CRUZ

Tutor: RUBÉN CUEVAS RUMÍN

EL TRIBUNAL

Presidente: _____

Vocal: _____

Secretario: _____

Realizado el acto de defensa y lectura del Proyecto Fin de Carrera el día ____ de _____
de 20__ en Leganés, en la Escuela Politécnica Superior de la Universidad Carlos III de
Madrid, acuerda otorgarle la CALIFICACIÓN de

VOCAL

SECRETARIO

PRESIDENTE

AGRADECIMIENTOS

Por fin ha llegado este momento. Ahora sí, después de tantos meses, puedo dar rienda suelta a mis emociones y escribir, sin preocuparme por el registro, la parte más personal de la memoria y agradecer el apoyo que he recibido.

Los primeros, mis padres, Juan y Cira. Los pilares fundamentales de mi vida. Espero que podáis estar, como mínimo, tan orgullosos de mí como lo estoy yo de vosotros. Gracias por apoyarme de forma incondicional siempre.

A mis hermanos, Juan y Alberto, porque son los mejores. Cuando aparecían los miedos, desánimos, las angustias,... ahí estabais vosotros. Muchas veces se han repetido en mi cabeza las palabras *“Tranquila Almudena, esto es una carrera de fondo”*. Gracias por protegerme y aguantarme.

A Rafa, mi compañero de viaje. No sé que habría sido de mí, si no te hubiera conocido... Soy muy afortunada por tenerte a mi lado. Gracias.

Mención especial para mi pequeño Noé, la alegría de la casa. No podía imaginar que una personita tan pequeña tuviese la capacidad de transmitir tanta energía positiva.

A mis amigos, por animarme y ayudarme a dar un enfoque distinto a las cosas, sacando el lado positivo.

A Rubén y Mika, mis tutores. Gracias por vuestro tiempo y por haberme ayudado en todos estos meses. Aunque ha costado mucho llegar hasta aquí, sin duda, ha merecido la pena.

RESUMEN DEL PROYECTO

Las Telecomunicaciones permiten la transmisión de información entre puntos distanciados. En ocasiones se crean, de forma intencionada, redes que reúnen distintos grupos de usuarios con un mismo fin: compartir información. Con este objetivo surgen las redes punto a punto.

BitTorrent es uno de los protocolos que se utilizan dentro de las redes punto a punto. Es útil para compartir distintos recursos en Internet: música, películas, series de televisión, programas, libros electrónicos y otros recursos multimedia. En la actualidad, esta y otras aplicaciones que permiten compartir recursos de interés general, reúnen un alto porcentaje del volumen de tráfico compartido a través de Internet.

En este Proyecto se presentan distintas herramientas, previamente desarrolladas en otros estudios, a través de las cuales se obtiene información relevante acerca de los recursos compartidos en el entorno de BitTorrent. En primer lugar se presenta *Crawler*, una herramienta que sienta las bases para poder rastrear información sobre los *.torrent* compartidos y los usuarios que publican estos recursos. En segundo lugar, siguiendo un orden cronológico dentro del desarrollo de las herramientas, se presenta *Torrent Guard*. Esta herramienta es capaz de identificar y clasificar el contenido compartido en el entorno de BitTorrent como malintencionado.

El objetivo principal de este Proyecto ha sido elaborar un portal web que reúne estas herramientas. De esta forma poder concentrar en un único punto las consultas sobre los principales publicadores de contenido en BitTorrent, así como sobre el estado de aquellos contenidos que se pretende descargar. A su vez, mostrar a los usuarios información relacionada sobre las investigaciones que se han llevado y se están llevando a cabo en el entorno de BitTorrent.

Palabras clave: punto a punto, p2p, BitTorrent, torrent, descarga, The Pirate Bay.

ABSTRACT

Telecommunications allows the transmission of information between points separated. Sometimes, intentionally, networks are created with the aim of grouping different users with one objective: share information. That is the objective of peer to peer networks.

BitTorrent is a protocol used by peer to peer networks. It is useful to share different kind of Internet resources: music, films, TV shows, software, e-book and other type of resources available. Nowadays this and other applications, which allow sharing resources, gather a high amount percentage of traffic volumen shared over the Internet.

In this Project different applications are presented, which have been developed in other studies, through relevant information about share resources is obtained inside BitTorrent enviroment. First of all *Crawler* is presented, this tool lays the foundations and tracks information about *.torrent* shared and all the users that publish this resources. Secondly, following a chronological order, *Torrent Guard* is presented. This tool is able to identify and classify the content shared in BitTorrent enviroment as malware.

The main objective in this Project has been developed an website in which gather together these applications. In this way all the searches about the main publishers and the status of *.torrent* that want to be downloaded, are collected in the same space. At the same time, show to users information about researches that have been done and are being carried out in BitTorrent enviroment.

Keywords: peer to peer, p2p, BitTorrent, torrent, download, The Pirate Bay.

ÍNDICE GENERAL

Agradecimientos	ii
Resumen del Proyecto	iii
Abstract.....	iv
Índice general.....	v
Índice de Figuras	ix
Índice de Tablas.....	xi
Capítulo 1	12
Introducción y objetivos	12
1.1 Introducción	12
1.2 Objetivos	13
1.3 Fases del desarrollo.....	14
1.4 Medios empleados.....	15
1.5 Estructura de la memoria	15
Capítulo 2	17
Redes P2P y protocolo BitTorrent	17
2.1 Introducción	17
2.2 Redes punto a punto	17
2.3 Protocolo BitTorrent.....	18
2.3.1 Introducción.....	18
2.3.2 Estructura de la red: principales agentes involucrados	19
2.3.3 Funcionamiento.....	19
2.3.4 Creación y publicación de un archivo .torrent.....	20
2.3.5 Descarga de un archivo .torrent	21
2.3.6 Motivación en el uso de BitTorrent.....	21
2.3.7 Clientes BitTorrent y páginas de descarga.....	22
2.4 Novedad: enlaces magnéticos	22

2.5	Problemática a resolver.....	23
Capítulo 3	25
MYPROBE: Monitoring, identifYing & PROfiling BitTorrent publishErs	25
3.1	Introducción.....	25
3.2	Punto de partida: Crawler.....	25
3.3	Analyzer.....	28
Capítulo 4	33
Torrent Guard	33
4.1	Introducción.....	33
4.2	Situación actual.....	33
4.3	Metodología de medida.....	34
4.4	Caracterización de los publicadores de contenido falso.....	35
4.4.1	Número y contribución de los publicadores malintencionados.....	35
4.4.2	Ubicación de los publicadores malintencionados.....	36
4.4.3	Utilización de las cuentas en The Pirate Bay.....	37
4.4.4	Estrategias de publicación.....	37
4.4.5	Estrategias para atraer a downloaders.....	38
4.5	Perfiles de los publicadores de contenido malintencionado.....	38
4.6	Torrent Guard.....	39
4.6.1	Implementación de Torrent Guard.....	40
4.6.2	Rendimiento de Torrent Guard.....	41
4.6.3	La eficiencia de Torrent Guard.....	42
Capítulo 5	43
Páginas web	43
5.1	Introducción.....	43
5.2	Punto de partida: Fake Detector.....	43
5.3	Estructura del portal web.....	44
5.3.1	Home.....	46
5.3.2	Myprobe: monitoring, identifying and profiling BitTorrent publishers.....	47
5.3.3	Torrent Guard.....	49
5.3.4	Related.....	51
5.3.5	People.....	53
5.3.6	Contact.....	54
5.4	HTML: HyperText Markup Language.....	56

5.5	Conflictos con navegadores: hojas de estilo en cascada CSS	56
5.6	Lenguaje de programación interpretado: PHP.....	57
Capítulo 6		58
Búsquedas dentro del portal web		58
6.1	Introducción	58
6.2	Consultas realizadas en MYPROBE	59
6.2.1	Publishers Information	59
6.2.2	TOP Publishers.....	64
6.2.3	TOP ISP.....	68
6.2.4	TOP Country.....	72
6.3	Consultas realizadas en Torrent Guard.....	75
6.3.1	Check your torrent.....	75
6.3.1.1	Opción 1: comprobar archivo .torrent.....	76
6.3.1.2	Opción 2: comprobar infohash.....	82
6.3.1.3	Opción 3: comprobar enlace magnético.....	86
Capítulo 7		93
Comportamiento del portal web en distintos navegadores.....		93
7.1	Introducción	93
7.2	Navegadores en el mercado.....	93
7.2.1	Internet Explorer	94
7.2.2	Mozilla Firefox.....	95
7.2.3	Google Chrome.....	95
7.2.4	Safari	95
7.3	Versiones de los navegadores utilizados.....	96
7.4	Resultados del portal web en distintos navegadores.....	97
Capítulo 8.....		101
Conclusiones.....		101
Capítulo 9.....		103
Presupuesto		103
9.1	Introducción	103
9.2	Presupuesto.....	103
9.2.1	Impacto económico.....	103
9.2.2	Impacto temporal	104
Glosario.....		106

Referencias.....	107
------------------	-----

ÍNDICE DE FIGURAS

Figura 1. Escenario BitTorrent [22].....	20
Figura 2. Aplicación Crawler [6].....	26
Figura 3. Diagramas de bloques completo de Crawler [6].....	28
Figura 4. Diagrama de flujo de Analyzer.....	31
Figura 5. Porcentaje de contenido malware en función del porcentaje de publicadores malintencionados [20].....	35
Figura 6. Principales ISP utilizados por usuarios de BitTorrent.....	36
Figura 7. Esquema de Torrent Guard [20].	40
Figura 8. Página de inicio para Fake Detector.	44
Figura 9. Estructura principal del portal web diseñado.....	45
Figura 10. Página principal.....	46
Figura 11. Apartado Related.	51
Figura 12. Subapartado Papers, enlaces disponibles.....	52
Figura 13. Apartado People.	53
Figura 14. Apartado Contact.	54
Figura 15. Mensaje de ayuda al usuario en el envío del formulario de contacto.	55
Figura 16. Confirmación de envío del mensaje.	55
Figura 17. Navegadores web.....	57
Figura 18. Acceso a Publishers Information.....	59
Figura 19. Mensaje de ayuda al usuario en Publishers Information.	60
Figura 20. Consulta de un usuario inventado en Publishers Information.	60
Figura 21. Resultado consulta de un usuario inventado en Publishers Information.	61
Figura 22. Resultado consulta de un usuario concreto (TvTeam) en Publishers Information.	62
Figura 23. Resultado consulta de un usuario concreto (digital_ripper) en Publishers Information.	63
Figura 24. Acceso a TOP Publishers.....	64
Figura 25. Mensaje de ayuda al usuario en TOP Publishers.	64
Figura 26. Mensaje de ayuda al usuario en TOP Publishers cuando se introducen valores no permitidos.	65
Figura 27. Consulta correcta en TOP Publishers.	65
Figura 28. Resultado consulta correcta en TOP Publishers.	66
Figura 29. Resultado del enlace More Information en TOP Publishers para digital_ripper.....	67
Figura 30. Mensaje de ayuda al usuario en TOP ISP.....	68

Figura 31. Mensaje de ayuda al usuario en TOP ISP cuando se introducen valores no permitidos.....	68
Figura 32. Consulta correcta en TOP ISP.....	69
Figura 33. Resultado consulta correcta en TOP ISP primera tabla.	70
Figura 34. Resultado consulta correcta en TOP ISP segunda tabla.....	71
Figura 35. Acceso a TOP Country.....	72
Figura 36. Resultado consulta correcta en TOP Country primera tabla.	73
Figura 37. Resultado consulta correcta en TOP Country segunda tabla.....	74
Figura 38. Acceso a Check your torrent.....	75
Figura 39. Mensaje de ayuda al usuario en la opción 1 de Check your torrent.....	76
Figura 40. Resultado consulta archivo enlaces.txt.....	77
Figura 41. Resultado consulta archivo M_Clan_-_Retrovision_(1995-2006)_[2006]_- _Pop.3496538.TPB.torrent.....	78
Figura 42. Resultado consulta archivo Law.and.Order.SVU.S14E06.HDTV.x264- LOL.7779386.TPB.....	79
Figura 43. Resultado consulta archivo Cubase_SX3_Magyar.oktat_oacute_- DVD.7639015.TPB.....	81
Figura 44. Mensajes de ayuda al usuario en la opción 2 de Check your torrent.	82
Figura 45. Resultado consulta infohash 65C59094DB4B206AE9B4FC95BB143005DA02BA82.....	83
Figura 46. Resultado consulta infohash bc1d09aecb189e069177cba1365961c3efe05d57.	84
Figura 47. Resultado consulta infohash 5453da50840c7ff7152c36d2c167c74f0cae706a.85	
Figura 48. Mensajes de ayuda al usuario en opción 3 de Check your torrent.	86
Figura 49. Resultado consulta enlace magnético probablemente correcto.	87
Figura 50. Resultado consulta enlace magnético desconocido.....	88
Figura 51. Resultado consulta enlace magnético correcto.	89
Figura 52. Resultado consulta enlace magnético correcto, sin información de usuario.....	91
Figura 53. Resultado consulta enlace magnético malicioso.....	92

ÍNDICE DE TABLAS

Tabla 1. Porcentaje de tráfico generado por BitTorrent en Europa [21].	18
Tabla 2. Principales navegadores de escritorio. Septiembre 2012.....	94
Tabla 3. Versiones de los navegadores utilizados.....	96
Tabla 4. Otros navegadores compatibles.....	96
Tabla 5. Características comunes en todas las páginas.....	98
Tabla 6. Características comunes en las páginas de consulta.	98
Tabla 7. Características comunes en las páginas de resultados.	99
Tabla 8. Características comunes en páginas con opción de descarga.....	99
Tabla 9. Características especiales en Contact.	99
Tabla 10. Porcentaje de fallos producidos en los distintos navegadores.....	100

CAPÍTULO 1

INTRODUCCIÓN Y OBJETIVOS

1.1 INTRODUCCIÓN

Las redes punto a punto [1] (también conocidas como *peer to peer* o *P2P*) son aquellas en las que todos los individuos se encuentran conectados unos con otros. De esta forma, todos se comunican con todos y actúan como iguales entre sí.

Una de las actividades más extendidas dentro de estas redes, es la transferencia de datos y compartición de recursos. A lo largo de los años, se han desarrollado distintos protocolos que definen y fijan las pautas para llevar a cabo estas actividades. BitTorrent es uno de estos protocolos. Define perfectamente el conjunto de normas y procedimientos que hay que realizar para conseguir la transferencia de contenidos. Siendo una de las aplicaciones que más volumen de datos mueve a través de Internet.

El concepto de BitTorrent también hace referencia al programa utilizado para llevar a cabo la descarga de contenido: cliente BitTorrent. En la actualidad existen diferentes clientes BitTorrent, como por ejemplo: BitTorrent [2] [3], Vuze, μ Torrent, Transmission, BitComet, etc.

El uso incremental de dichas aplicaciones hace que los recursos que se comparten aumenten. La problemática surge cuando también se incrementan los recursos malintencionados.

Este Proyecto está orientado a elaborar un portal web de consulta que integra el acceso a diferentes herramientas, como: *Crawler* y *Torrent Guard*, las cuales han sido previamente desarrolladas en otros proyectos. A través de dichas aplicaciones se puede extraer información

de archivos *.torrent* compartidos dentro del entorno de BitTorrent (por ejemplo: a través del portal *The Pirate Bay* [4]).

La herramienta principal con la que se recaba información es: Crawler. Este programa monitoriza y consigue información sobre los archivos *.torrent* publicados en The Pirate Bay gracias al formato de distribución RSS⁽¹⁾. El portal web accede a toda la información obtenida a través del apartado *MYPROBE*.

En paralelo, analizando la información obtenida, se genera otra base de datos donde se guardan las referencias a los *.torrent* dañados. Finalmente, en el conjunto de páginas web se muestran de una forma atractiva y visual los resultados encontrados con ayuda de las aplicaciones mencionadas. Además, en el portal, se añaden referencias a documentos publicados por las personas que han participado en este Proyecto, englobando así gran parte del trabajo realizado.

Es importante conocer si un archivo que se quiere descargar está en buen estado o por el contrario es un elemento dañino para el usuario. A través del portal web, el usuario puede comprobar el estado del archivo que desee descargar gracias a Torrent Guard. Este hecho, permite analizar la intencionalidad de los posibles elementos infectados en páginas como The Pirate Bay.

Por otro lado, en las distintas páginas web muestra información sobre los usuarios que comparte información en la red. De esta forma conocer si dicho usuario es de confianza o por el contrario el interés del mismo es compartir archivos dañados.

En este capítulo se van a exponer los objetivos principales perseguidos con este Proyecto. A su vez se explicarán las diferentes fases de desarrollo y se enumerarán los medios empleados para llevar a cabo el desarrollo de las distintas aplicaciones. Finalmente, se resumirán los siguientes capítulos que componen esta memoria.

1.2 OBJETIVOS

El objetivo principal del presente Proyecto es elaborar un portal web a través de la cual un usuario puede comprobar la veracidad de un archivo *.torrent* que desee descargar.

Un usuario puede publicar y compartir información a través de portales como The Pirate Bay [5]. Este portal actúa como motor de búsqueda y servidor para todo tipo de archivo multimedia.

En infinidad de ocasiones el archivo descargado no coincide con aquello que el usuario realmente quiere descargar. Por este motivo, es interesante poder verificar previamente si el archivo objeto de descarga es correcto.

⁽¹⁾ RSS: en inglés *Really Simple Syndication* [33], formato *.XML* utilizado para difundir información actualizada frecuentemente del contenido de una página.

Con ayuda de la aplicación ya desarrollada Torrent Guard, se puede advertir al usuario de la falsedad de los archivos a través del portal web. Además, gracias a la monitorización realizada con Crawler y los patrones de detección utilizados, la identificación de archivos multimedia falsos puede realizarse más rápidamente que en el propio portal de descargas.

Además, el usuario podrá consultar más información relacionada con los usuarios que publican contenidos en The Pirate Bay. Conocer el número de archivos compartidos en el portal por dichos usuarios, dentro de qué categoría publican mas *.torrent*, los ISP⁽²⁾ utilizados, incluso el país o ciudad de donde procede la descarga.

Independientemente del usuario identificado en The Pirate Bay, a través del portal web elaborado, se pueden conocer otros datos de interés y de gran valor para los usuarios que comparten contenidos multimedia. Por ejemplo: los listados con los principales usuarios que publican contenidos, los principales servidores de Internet utilizados,... así como filtrar esta información en función de una categoría concreta (*video*, música, juegos,...).

Este Proyecto comparte las mismas líneas de investigación que otros proyectos desarrollados por el Departamento de Ingeniería Telemática de la Universidad Carlos III. Por este motivo, en el portal web desarrollado, se mencionan otros trabajos de investigación con los que este Proyecto está ligado.

1.3 FASES DEL DESARROLLO

En primer lugar, se realiza la primera toma de contacto con los elementos que estaban desarrollados. Estos serían: Crawler, aplicación de consulta elaborada por David Rubio [6] y *plugin*⁽³⁾ para Vuze realizado por Beatriz Calvo [27].

Al mismo tiempo, comprender y analizar estas aplicaciones. En concreto, interpretar el código de la aplicación Crawler para entender cómo funciona en detalle e interactúa con la página The Pirate Bay. También conocer como se generan y consultan las bases de datos generadas.

Definir las pautas de detección que se van a utilizar para encontrar los archivos malintencionados publicados y que son compartidos a través de portales BitTorrent como, por ejemplo: The Pirate Bay.

Generar dicha base de datos con todos los archivos pautados y destacados como corruptos. Recopilar toda la información útil para su detección y por la que se seleccionan estos archivos como malintencionados.

Establecer la información básica que se quieren mostrar en el portal web a desarrollar. Estructurarla de forma modular para que resulte cómodo modificar las diferentes secciones.

⁽²⁾ ISP: en inglés *Internet Service Provider* [34], sirve para conectar a los usuarios a Internet a través de diferentes tecnologías.

⁽³⁾ Plugin: complemento o aplicación.

Diseñar el estilo y la apariencia del portal web de forma que resulte visualmente atractivo, útil y sencillo de utilizar. Compatibilizar su funcionalidad con los principales navegadores web: Internet Explorer, Mozilla Firefox, Google Chrome o Safari.

Verificar que para todos los escenarios posibles dentro del portal web, éste funciona perfectamente y que su comportamiento es preciso en cada una de las búsquedas.

Finalmente, elaborar la memoria del Proyecto Fin de Carrera en la que se resume el trabajo realizado.

1.4 MEDIOS EMPLEADOS

La mayor parte del trabajo se ha realizado de forma presencial en el laboratorio de la Universidad Carlos III. La ubicación del laboratorio es 4.1.F04.

El equipo utilizado tiene las siguientes características:

- Sistema operativo: Ubuntu.
- Versión: 10.10 (maverick)
- Núcleo Linux 2.6.35-25-generic-pae
- GNOME 2.32.0
- 4 Procesadores Intel® Core™2 Quad CPU q9400 @ 2.66GHz

Dentro de la fase inicial del Proyecto, se ha utilizado Eclipse SDK versión 3.5.0 como entorno de desarrollo.

Para programar el grueso del portal web se ha utilizado un editor de texto normal incluido dentro de las aplicaciones del propio ordenador.

Se han instalado distintos navegadores web para comprobar el comportamiento del portal web, ya que las hojas de estilo definidas se interpretan de forma distinta en función del navegador. Se ha optado por comprobar este comportamiento sobre: Internet Explorer, Mozilla Firefox, Google Chrome y Safari.

1.5 ESTRUCTURA DE LA MEMORIA

En este apartado se resumen brevemente cada uno de los capítulos que forman esta memoria:

- **Capítulo 1: Introducción y objetivos.**
Este primer capítulo sirve como presentación de las redes punto a punto y el vínculo existente con el protocolo BitTorrent. A su vez, se justifica la realización de este Proyecto y se exponen los objetivos que pretende lograr.

- **Capítulo 2: Redes P2P y protocolo BitTorrent.**

En este capítulo se presentan las redes punto a punto y se explica, de forma detallada, en qué consiste el protocolo de descarga BitTorrent. Entre otros conceptos, se describe: su funcionamiento y los agentes involucrados.

- **Capítulo 3: MYPROBE: Monitoring, identifyIng & PROfiling BitTorrent publishErs.**

Resume el punto de partida del Proyecto. Presenta la aplicación *Crawler* repasando sus puntos principales. Detalla la aplicación *Analyzer*: funcionamiento y los objetivos que persigue.

- **Capítulo 4: Torrent Guard.**

Capítulo basado en una de las investigaciones previas realizadas por el Departamento de Ingeniería Telemática de la Universidad Carlos III. Se detallan los distintos escenarios dentro de la publicación de contenidos en BitTorrent. Se identifican los principales perfiles de publicadores malintencionados. Se explica cómo se ha obtenido la información mostrada en el apartado *Torrent Guard*, dentro del portal web desarrollado.

- **Capítulo 5: Páginas web.**

Exposición de cómo se organizan las aplicaciones desarrolladas dentro del portal web. Explicación de los distintos apartados y subapartados que se pueden encontrar.

- **Capítulo 6: Búsquedas dentro del portal web.**

Se muestran diferentes ejemplos de consultas realizadas sobre el portal web.

- **Capítulo 7: Comportamiento del portal web en distintos navegadores.**

Se realiza una comparación y análisis de cómo se comporta el portal web en diferentes navegadores.

- **Capítulo 8: Conclusiones.**

Incluye las conclusiones extraídas tras el trabajo realizado.

- **Capítulo 9: Presupuesto.**

Último capítulo de la memoria que analiza el coste económico del Proyecto.

CAPÍTULO 2

REDES P2P Y PROTOCOLO BITTORRENT

2.1 INTRODUCCIÓN

El objetivo de este capítulo es introducir el concepto de redes punto a punto y describir en detalle el protocolo BitTorrent. Presentar las características principales, el funcionamiento o la estructura del protocolo, centrarán el grueso de esta sección. Además se abordará tanto la motivación de este tipo de redes, como su uso en la actualidad.

Para cerrar este capítulo se expondrán algunas debilidades que presenta BitTorrent y que se pretenden resolver, o al menos mitigar, a través de las herramientas desarrolladas.

2.2 REDES PUNTO A PUNTO

Las redes punto a punto [1] son redes de distribución en las que todos los elementos actúan como iguales entre sí, es decir, no existen jerarquías entre elementos dentro de la misma red. Un elemento puede actuar como cliente y/o servidor. La principal ventaja de las redes punto a punto reside en la gran facilidad a la hora de compartir recursos.

El principio de funcionamiento de dichas redes se basa en que los usuarios comienzan a compartir recursos desde el instante inicial en el que comienzan a descargar algún tipo de

archivo. Así, un usuario que tiene parte de un contenido colabora favoreciendo la descarga del mismo al compartirlo con otros.

Existe un amplio abanico de recursos que pueden ser compartidos a través de Internet: música, películas, aplicaciones, software, juegos, libros electrónicos, series... Dichos recursos son especialmente demandados por los usuarios, de ahí que este tipo de redes tengan tanto auge.

2.3 PROTOCOLO BITTORRENT

2.3.1 INTRODUCCIÓN

BitTorrent es un protocolo [2] diseñado para el intercambio de archivos punto a punto por Bram Cohen. Es una manera de transferir archivos de cualquier tamaño de forma rápida y eficiente. En la actualidad esta aplicación es usada por miles de usuarios y responsable de gran parte del tráfico que se comparte en Internet.

Algunas publicaciones recientes [21] reflejan que un alto porcentaje del tráfico de datos de Internet está dedicado a la descarga de contenido utilizando el protocolo BitTorrent. Como se observa en la siguiente tabla, según datos del último estudio de Sandvine, BitTorrent agrupa más del 31% del tráfico de subida en redes europeas y un 17% del tráfico de bajada.

Europe Fixed						
Rank	Upstream		Downstream		Aggregate	
	Application	Share	Application	Share	Application	Share
1	BitTorrent	31.69%	HTTP	19.46%	BitTorrent	20.32%
2	eDonkey	18.23%	YouTube	18.00%	HTTP	17.70%
3	HTTP	11.29%	BitTorrent	17.20%	YouTube	15.25%
4	YouTube	5.24%	eDonkey	6.96%	eDonkey	9.39%
5	Skype	2.48%	Flash Video	5.62%	Flash Video	4.70%
6	SSL	2.47%	RTMP	2.80%	RTMP	2.47%
7	Teredo	2.31%	Facebook	2.54%	Facebook	2.43%
8	Facebook	2.02%	MPEG Streaming	1.98%	SSL	1.74%
9	Flash Video	1.34%	iTunes	1.68%	MPEG Streaming	1.66%
10	BBC iPlayer	1.27%	SSL	1.54%	iTunes	1.53%
	Top 10	78.34%	Top 10	77.78%	Top 10	77.19%

SOURCE: SANDVINE NETWORK DEMOGRAPHICS




Tabla 1. Porcentaje de tráfico generado por BitTorrent en Europa [21].

El nombre de BitTorrent también hace referencia al cliente BitTorrent, es decir, la aplicación capaz de ejecutar el protocolo. En la actualidad existen distintos clientes que pueden ser

utilizados como, por ejemplo: Vuze, μ Torrent, BitTorrent, Transmission o BitComet. Se pueden encontrar distintos tutoriales en Internet que facilitan su uso.

Un usuario que participa en la transferencia de archivos actúa como fuente así como peticionario del recurso compartido. Es decir, se descarga parte del contenido de otros usuarios y a la vez comparte el contenido que ya tiene. Cuantos más usuarios participen en la transferencia del archivo, más rápidamente se realiza la descarga.

2.3.2 ESTRUCTURA DE LA RED: PRINCIPALES AGENTES INVOLUCRADOS

Los principales agentes que participan dentro de BitTorrent son:

- **Seeder:** usuario inicial que publica y comparte un archivo *.torrent* y que posee una copia completa del mismo.
- **Leecher:** usuario que descarga y comparte un archivo *.torrent* pero que no tienen una copia completa.
- **Peer:** usuario que participa en el intercambio de archivos *.torrent*.
- **Tracker:** servidor que realiza el seguimiento de seeders y leechers. No posee una copia de los archivos que están siendo compartidos. Su objetivo es gestionar el proceso de transferencia de archivos.
- **Swarm o enjambre:** se entiende como el grupo de seeders y leechers que están compartiendo el mismo *.torrent*.

2.3.3 FUNCIONAMIENTO

La idea de Bram Cohen [2] fue fragmentar el archivo que quería transferir en pequeños trozos. Así, el archivo es descargado trozo a trozo a través de la conexión a distintos usuarios que actúan como fuente, para poder obtener el archivo completo. Es eficiente porque se realizan descargas más rápidas.

El tamaño [22] que tienen estos k trozos es el mismo y típicamente es 256 KB o 4 MB. A su vez, estos trozos se dividen en bloques con un tamaño de 16 KB que son las unidades intercambiadas entre los peers.

El factor [2] más importante para determinar si un enjambre puede continuar permitiendo a los peers completar la descarga de un *.torrent* es la disponibilidad. La disponibilidad de un *.torrent* es el número de copias completas del *.torrent* que están distribuidas en la red en la que se está conectado, incluyendo los trozos ya adquiridos.

Para poder localizar cada agente, se necesita centralizar la localización de los peers con la información de su dirección IP. El Tracker de BitTorrent es un servidor que centraliza esta información. Simplemente necesita la dirección IP y el puerto a través de los cuales los seeders se conectan al enjambre para compartir.

Dada la naturaleza de BitTorrent, las velocidades de descarga no están garantizadas para ningún *.torrent* dentro del enjambre. Mientras que en un enjambre pueden obtenerse velocidades muy altas, en otros no. Esto se debe al hecho de que BitTorrent es un protocolo punto a punto, por lo tanto, depende de las velocidades de subida de los peers a los que se está conectado para calcular la velocidad de descarga disponible.

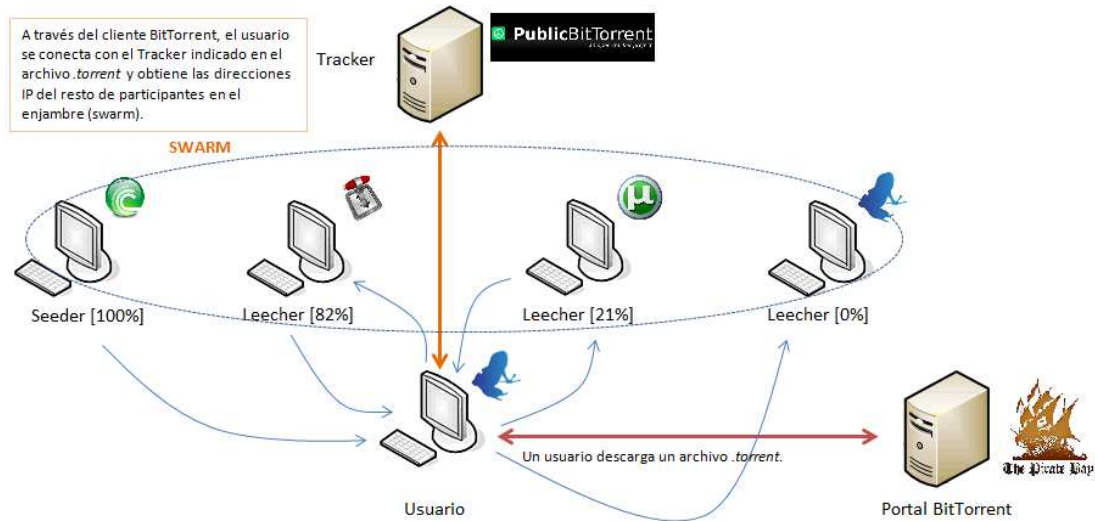


Figura 1. Escenario BitTorrent [22].

2.3.4 CREACIÓN Y PUBLICACIÓN DE UN ARCHIVO .TORRENT

Cuando un usuario quiere publicar algún contenido dentro de BitTorrent [20], [22], en primer lugar debe crear un archivo *.torrent*. Este archivo incluye, junto con otra información, la dirección IP (o direcciones) del tracker (o trackers) que gestionan el proceso de descarga del contenido.

Después, este archivo se publica en uno o varios portales de BitTorrent que actúan como puntos de encuentro para poder centralizar la publicación y descarga de estos archivos *.torrent*. Para ello, los usuarios que quieren publicar un contenido usan una cuenta de usuario con un identificador específico (*username*), creado en dicho portal. Además, este publicador distribuye la primera copia del contenido actuando como seeder inicial en el enjambre en el que está asociado. Por lo tanto, el contenido publicado, puede ser identificado gracias a la dirección IP del seeder inicial que distribuye este contenido y por el nombre de usuario utilizado dentro del portal de BitTorrent.

Algunos portales de BitTorrent, como The Pirate Bay, indexan millones de *torrents* y reciben millones de visitas. Además, ofrecen información sobre cada uno de los archivos *.torrent* que tienen: nombre, fecha de publicación, categoría del *.torrent*, número de archivos asociados,

tamaño del archivo completo, nombre de usuario que lo ha publicado, número de seeders y leechers que participan en el enjambre y una descripción algo más detallada del contenido.

2.3.5 *DESCARGA DE UN ARCHIVO .TORRENT*

Como se observa en la Figura 1, cuando un usuario desea descargar [20], [22], algún archivo multimedia, en primer lugar descarga el archivo *.torrent* asociado al contenido que desea. Éste estará publicado dentro de un portal de BitTorrent como, por ejemplo: The Pirate Bay. El archivo puede abrirse mediante un cliente de BitTorrent como, por ejemplo, Vuze.

A través del cliente BitTorrent, el usuario se conecta a la dirección IP de uno de los trackers especificados en el archivo *.torrent*. El tracker, como se ha comentado anteriormente, es el servidor central que conoce las direcciones IP de cada uno de los peers que están compartiendo dicho archivo. Por tanto, este servidor, proporciona al usuario un listado con direcciones IP (entre 50 y 200) que participan en dicho enjambre, junto con el número de seeders y leechers que lo forman. Finalmente, el usuario comienza a descargar el contenido de las direcciones IP que ha obtenido.

2.3.6 *MOTIVACIÓN EN EL USO DE BITTORRENT*

En la actualidad BitTorrent es una de las aplicaciones más utilizadas en Internet. Es usada a diario por miles de usuarios, siendo responsable de gran parte del tráfico de datos. Como se desprende del estudio realizado por Sandvine [21], sólo en Europa es responsable de más del 31% del tráfico de subida y más del 17% del tráfico de bajada. Se estaría hablando de que es responsable de 1 de cada 5 GB transferidos por las redes fijas.

Si se agrupan los datos de tráfico para las aplicaciones de descarga punto a punto que aparecen en la Tabla 1, se observa que suponen más de un tercio del volumen total de los datos compartidos a través de la red.

En los últimos meses, tras el cierre de algunos espacios web de descarga de archivos o *streaming*⁽⁴⁾, como son Megaupload o Megavideo, el volumen de tráfico en las redes punto a punto se ha visto favorecido, incrementando su tráfico. Los usuarios buscan poder compartir entretenimiento multimedia como, por ejemplo: películas, música, series, juegos, programas,... El valor añadido es que, a través de BitTorrent, el coste es nulo.

⁽⁴⁾ Streaming: descarga en tiempo real.

2.3.7 CLIENTES BITTORRENT Y PÁGINAS DE DESCARGA

Para gestionar una descarga utilizando BitTorrent, como se ha explicado anteriormente, es necesario buscar el archivo *.torrent* en algunos de los portales disponibles y, además, disponer de un cliente BitTorrent capaz de ejecutar la descarga.

En la actualidad existen diferentes clientes capaces de interpretar el protocolo en el ordenador de un usuario. Estos programas se pueden descargar fácilmente desde cualquier página de descarga de software. Algunos ejemplos de clientes BitTorrent son [11]: BitTorrent, μ Torrent, Vuze, Transmission o BitComet. Todos ellos cuentan con características puntuales que los distinguen entre ellos. Sin embargo, todos se encargan de gestionar la conexión para que los usuarios puedan completar sus descargas.

La búsqueda de archivos *.torrent* se puede llevar a cabo a través de distintos portales BitTorrent. A diferencia de otros protocolos descarga punto a punto, son los usuarios los encargados de encontrar el contenido concreto que desean descargar. Hay ciertas páginas dedicadas a la publicación de archivos *.torrent*, las principales web de descarga de torrents son: The Pirate Bay, Mininova, IsoHunt o Torrentz.

2.4 NOVEDAD: ENLACES MAGNÉTICOS

Los enlaces magnéticos [29] son un tipo de enlace utilizado para identificar un contenido sin tener que especificar su nombre o ubicación, sino solo uno o más valores *hash* obtenidos a partir de funciones hash criptográficas. De una forma más técnica se puede decir que un enlace magnético es un esquema URI⁽⁵⁾ para definir hipervínculos que normalmente usan una URN⁽⁶⁾ para enlazar, ya que hace referencia a un archivo en base a su contenido o metadato, no en base a su ubicación como hacen las URL.

En la actualidad, The Pirate Bay ha añadido este tipo de enlaces a los archivos publicados en su portal. De esta manera, existe la oportunidad de descargar el archivo *.torrent* o acceder al enlace magnético, dando prioridad a esta segunda opción y estando ya en decadencia la primera. No significa que los archivos *.torrent* hayan desaparecido, simplemente implica que se está apostando por esta nueva forma de publicar archivos.

La principal ventaja [32] que motiva el uso de enlaces magnéticos es que los archivos *.torrent* no tienen que estar almacenados en los servidores de los portales de descarga. El propio enlace contiene toda la información necesaria para descargar el archivo.

La estructura de estos enlaces [32] en el caso de BitTorrent, sería la siguiente:

⁽⁵⁾ Esquema URI: [30] forma organizada de estructurar de forma uniforme los identificadores de recursos.

⁽⁶⁾ URN: [31] en inglés Uniform Resource Name, sirven para identificar el recurso web pero sin indicar exactamente dónde se encuentra el objeto.

- **magnet:?**: es la secuencia que identifica el comienzo del enlace en cuestión. Es común para todos los enlaces.
- **xt=urn:**: sirve para identificar la red. Por ejemplo, en el caso de BitTorrent, la secuencia completa sería: `xt=urn:bthi`. Donde `bthi` significa: BitTorrent *hash information*.
- **Información hash**: a continuación de la secuencia anterior iría la información hash del archivo que se quiere compartir.
- **dn**: sirve para indicar el nombre del archivo.
- **tr**: sirve para indicar un rastreador para las descargas en BitTorrent.

De esta manera, un ejemplo de enlace magnético podría ser: **magnet:?xt=urn:btih:fdec0abe886123999baebf3164bdcdd74503587e&dn=Maroon+5+-+Moves+Like+Jagger&tr=udp%3A%2F%2Ftracker.openbittorrent.com%3A80&tr=udp%3A%2F%2Ftracker.publicbt.com%3A80&tr=udp%3A%2F%2Ftracker.istole.it%3A6969&tr=udp%3A%2F%2Ftracker.ccc.de%3A80**.

2.5 PROBLEMÁTICA A RESOLVER

Como se ha comentado anteriormente, BitTorrent es un bien común en la actualidad de las Telecomunicaciones. No cabe duda que, con el paso de los años, se ha seguido explotando su funcionalidad principal: compartir recursos.

Algunas investigaciones [20] analizan la vulnerabilidad del protocolo BitTorrent, mientras que otras han centrado su interés en analizar la falta de privacidad del mismo. Dedicando menos atención a examinar aquellos aspectos relacionados con la seguridad.

Como se ha revelado en otros estudios [24], los distintos escenarios de BitTorrent están sufriendo una constante contaminación en sus contenidos. Según datos extraídos de estos estudios, se hablaría de que el 30% de los contenidos compartidos en BitTorrent estarían asociados a contenidos falsos o *malware*.

Este problema supone que si un usuario descarga contenido de un portal de BitTorrent, como por ejemplo The Pirate Bay, podría suceder que:

- el contenido descargado no corresponda con lo esperado. Por ejemplo: que se descargue una película cuyo título no coincide con lo esperado.
- el contenido sea malintencionado. Por ejemplo: que al ejecutar el contenido el ordenador se estropee.
- el archivo sea erróneo. Por ejemplo: no se permita ejecutar el archivo en el ordenador una vez descargado.
- la descarga forme parte de alguna estafa. Por ejemplo: la ejecución del archivo descargado redirija a páginas web externas, en las cuales se pidan datos personales para poder ejecutar el archivo.

- el contenido sea una versión falsa de forma intencionada. Por ejemplo: las propias agencias de antipiratería publican versiones falsas de los contenidos con derecho de autor que quieren proteger.
- la descarga se produzca de forma correcta.

Con el objetivo de continuar dentro de las mismas líneas de investigación en el ámbito de aplicaciones como BitTorrent, se plantea enfocar el análisis desde el punto de vista de la seguridad. De esta forma se propone una solución práctica para mitigar la presencia de contenidos falsos dentro de BitTorrent.

Torrent Guard identifica a los publicadores de contenido falso por su dirección IP. Esta herramienta, desarrollada con anterioridad en otro proyecto, está incluida dentro del portal web desarrollado en este Proyecto.

Por un lado, permite automatizar y acelerar el proceso de detección, incluso antes que el propio portal de BitTorrent. Por otro lado, aporta una solución válida independientemente del portal de descarga del archivo.

CAPÍTULO 3

MYPROBE: MONITORING, IDENTIFYING & PROFILING BITTORRENT PUBLISHERS

3.1 INTRODUCCIÓN

En este capítulo se describe el punto de partida desde el cual se comienza a trabajar en este Proyecto. El capítulo empieza describiendo los aspectos fundamentales de la aplicación Crawler y la actividad que realiza.

A continuación, se describe el funcionamiento de Analyzer, una aplicación que nace a partir de Crawler. Desarrollada con el fin de obtener un listado detallado de aquellas direcciones IP malintencionadas. Las cuales introducen contenidos falsos dentro de los portales de BitTorrent.

3.2 PUNTO DE PARTIDA: CRAWLER

La aplicación de referencia a partir de la cual se comienza a trabajar se llama Crawler [6]. El objetivo principal de esta aplicación es averiguar quién sube los ficheros que se publican en BitTorrent. Para desarrollar esta actividad, la aplicación monitoriza de forma continua el portal web The Pirate Bay.

Crawler recopila toda la información relativa de aquellos publicadores de contenido dentro BitTorrent. De esta forma, obtiene información de los archivos *.torrent* monitorizados y de los usuarios que los publican. La Figura 2 muestra el aspecto de la aplicación antes de iniciar sesión en ella.


The image shows a web browser window displaying a login page titled "LOGIN CRAWLER". The page has a light blue header and a white main content area. Below the title, there are two input fields: "User" and "Password", each with a small icon to its left. Below these fields is a "Login" button. At the bottom of the page, it says "Powered by David Rubio".

Figura 2. Aplicación Crawler [6].

El programa monitor está compuesto por distintos bloques, cada uno de ellos encargado de ejecutar acciones concretas que se describen a continuación:

- **Módulo RSS:** este módulo se ejecuta de forma ininterrumpida con el objetivo de obtener nuevos torrents que se hayan publicado en The Pirate Bay. Gracias al *feed* RSS⁽⁷⁾ de este portal de BitTorrent, se obtiene un archivo *.xml* donde aparecen estos torrents. La información se procesa y se almacena en la base de datos correspondiente.
- **Módulo de conexión al tracker:** aquí se llevan a cabo dos conexiones. Por un lado, la primera conexión tiene el objetivo de conseguir información complementaria relacionada con los torrents descargados. Por tanto, en esta primera conexión el tracker devuelve información de los seeders y leechers, así como el *.torrent* en cuestión. Por otro lado, la segunda conexión, sirve para obtener las direcciones IP de los peers indicados anteriormente.
- **Módulo de conexión a los peers:** tras obtener el listado del módulo anterior, esta parte de Crawler se dedica a identificar los peers. Para ello, se conecta a las direcciones IP de los peers indicados en el módulo anterior y analiza uno a uno el mensaje *bitfield*⁽⁸⁾ recibido. Además, una vez etiquetado el peer y ayudándose de otras herramientas (GeoIP proporcionada por MaxMind⁽⁹⁾) identifica el ISP, país y ciudad. Toda la información recabada se almacena en las bases de datos correspondientes y se guarda.

⁽⁷⁾ Feed RSS: es un servicio que ofrecen algunas páginas web. Su objetivo es mantener informados a los usuarios que estén suscritos al servicio y que reciban las últimas novedades publicadas en la página.

⁽⁸⁾ Bitfield: mensaje intercambiado por los peers, que indica el número de partes del fichero que posee el usuario al que pregunta.

⁽⁹⁾ MaxMind: empresa que proporciona herramientas de geolocalización.

- **Bases de datos:** Crawler utiliza una base de datos MySQL [9] con el objetivo de almacenar toda la información de los torrents descargados. El proceso de inserción, actualización y borrado se puede ejecutar de manera rápida, permitiendo mayor eficiencia en las consultas que se realizan. A continuación se describen brevemente las tablas utilizadas:
 - *Tabla TORRENT:* almacena toda la información extraída de los *.torrent* monitorizados por Crawler.
 - *Tabla PEERS:* almacena información de los usuarios, ya sean seeders o leechers que publican o descargan torrents en The Pirate Bay.
 - *Tabla FILES:* almacena los torrents que han sido analizados para evitar duplicidades en las tablas y en las consultas.
 - *Tabla UPLOADERS:* se almacena información manualmente sobre los usuarios.

La unión de estos módulos hace posible obtener, almacenar y analizar información acerca de los torrents publicados en The Pirate Bay. A continuación se describe el funcionamiento completo de la aplicación:

- En primer lugar, la aplicación comienza a ejecutarse a través del módulo RSS. Como se ha mencionado anteriormente, este modulo obtiene una lista con los últimos torrents publicados en The Pirate Bay.
- A continuación, se descargan los torrents obtenidos con el fin de obtener más información acerca de ellos y poder iniciar las conexiones necesarias con el/los tracker/s.
- Después, con la lista completa de peers, se realizan las conexiones a cada uno de ellos para poder averiguar cuál tiene el archivo completo.
- Al haber obtenido un *.torrent* momentos después de su publicación, los usuarios que dispongan del fichero completo serán, con casi toda seguridad, los responsables de haber subido el contenido a la red, es decir, seeders.
- En las diferentes tablas se van almacenando todos los datos recopilados de los torrents analizados durante el proceso de monitorización.

El siguiente diagrama muestra de manera esquemática los principales componentes del sistema desarrollado y la relación existente entre cada uno de ellos.

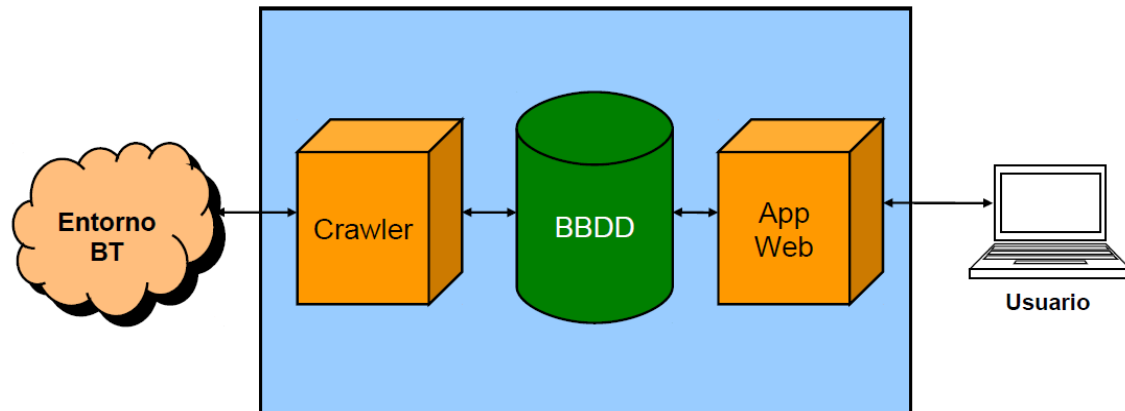


Figura 3. Diagramas de bloques completo de Crawler [6].

Toda la información recopilada acerca de los torrents que monitoriza Crawler, puede consultarse a través de su propia aplicación web. El enlace a esta aplicación es <http://163.117.140.136/web/login.php>. Por tanto, un usuario podrá conocer más acerca de los publicadores de contenido y los torrents que deseen descargar. Para consultar esta aplicación es necesario autenticarse y rellenar un formulario de consulta.

3.3 ANALYZER

El trabajo comienza analizando el código fuente de la aplicación anteriormente descrita. Para ello se ejecuta la aplicación utilizando el programa Eclipse [7], [8]. Una vez instalados los distintos módulos de la aplicación, se revisan los diferentes archivos implicados en Crawler.

El objetivo principal en esta parte del Proyecto ha sido entender los detalles que hacen posible la ejecución de la aplicación Crawler. Así poder conocer las particularidades del código y comenzar a progresar en nuevos desarrollos.

A partir de aquí se comienza a trabajar en una nueva aplicación o evolución del código, que extraiga otro tipo de información sobre los torrents monitorizados. Esta aplicación recibe el nombre de Analyzer. Su objetivo es generar un listado con todas las direcciones IP que se pueden clasificar como malintencionadas.

Analyzer presta atención y analiza la información recabada dentro de los ficheros: *torrent.data* y *seeder.data*, ambos generados por Crawler.

Cada línea dentro de *torrent.data* tiene el siguiente formato:

id_torrent	pub_time	category	subcategory	username	name	length
------------	----------	----------	-------------	----------	------	--------

- **id_torrent:** número secuencial que identifica el *.torrent* dentro de The Pirate Bay.
- **pub_time:** fecha en la cual se ha publicado el *.torrent* dentro del portal de BitTorrent.

- **category:** tipo de contenido. Puede ser: audio, video, juegos, porno, aplicaciones u otros.
- **subcategory:** dentro de las categorías anteriores, cabe la posibilidad de clasificar el contenido en subcategorías. Por ejemplo: música, audio libros, películas, series, aplicaciones para Android, Windows, MAC u otros, juegos para ordenadores, consolas o dispositivos móviles, comics, fotos u otros.
- **username:** nombre que utiliza un usuario para publicar contenido en The Pirate Bay.
- **name:** título del *.torrent* publicado.
- **length:** dimensión del archivo en *bytes*.

Por su parte, los campos del fichero *seeder.data* son:

id_torrent	pub_time	seeder	leecher	ip
Port	mask24 / mask16	ISP	country	city

- **id_torrent:** número que identifica el *.torrent* dentro de The Pirate Bay.
- **pub_time:** fecha en la cual se ha publicado el *.torrent* dentro del portal de BitTorrent.
- **seeder:** número de usuarios que poseen una copia completa del archivo.
- **leecher:** número de peers que tienen parte del *.torrent*.
- **ip:** dirección IP del usuario que ha publicado el *.torrent*.
- **port:** puerto.
- **mask24 / mask16:** máscara de red de la dirección IP.
- **ISP:** proveedor de servicio de Internet utilizado para publicar el *.torrent*.
- **country:** país desde el que se publica el contenido.
- **city:** ciudad desde la que se publica el contenido.

El funcionamiento de esta aplicación se representa en el diagrama de flujo de la Figura 4. A continuación se describe brevemente cómo se ejecuta esta aplicación:

- Se arranca la aplicación Crawler. A través de *BTCrawler.java*, que es su clase principal, se crea una instancia de Analyzer que llama al método *run()* de la misma.
- Se inicializan las variables necesarias. Los archivos que genera Analyzer se almacenan dentro de una carpeta llamada Analyzer. Estos archivos son:
 - *copyTorrents.data*: es una copia del fichero *torrent.data*
 - *copySeeders.data*: es una copia del fichero *seeder.data*
 - *fakes.data*: contiene el listado de posibles direcciones IP malintencionadas
 - *IPFakes.data*: contiene un listado con las direcciones IP identificadas como maliciosas.

Por este motivo, en primer lugar se comprueba qué archivos existen en esta carpeta y se borran aquellos que no interesan.
- Para poder seguir el proceso de análisis se van mostrando por pantalla mensajes de seguimiento. Así se conoce en qué punto de ejecución se encuentra el programa.

- En este punto se realiza una copia de los archivos mencionados anteriormente: *torrent.data* y *seeder.data* obteniendo sus homólogos en Analyzer. Así, no existirá conflicto de lectura o escritura si el programa principal los modifica.
- Si es la primera vez que se ejecuta el programa, se inicializan los valores de las variables con la información encontrada en *fakes.data* e *IPFakes.data*.
- Se comienza a leer línea a línea la información de *copyTorrents.data*. Fijando el límite de búsqueda en la antigüedad de los torrents almacenados, se comienza extrayendo el nombre de usuario (username) de cada línea. Si dicho *.torrent* no está almacenado en el listado de *.torrent* mirados previamente, ni el usuario ha sido consultado con anterioridad o se encuentra almacenado en *fakes.data*, entonces se consulta la página The Pirate Bay.
- Con el identificador del *.torrent* y el nombre de usuario, se consulta en el portal de BitTorrent para averiguar si el usuario ha sido borrado de la página The Pirate Bay. Se almacena el identificador para no volver a preguntar por él y evitar así duplicidades.
- Si el usuario ha sido borrado del portal de BitTorrent, se busca dentro de *copySeeder.data* el identificador del *.torrent*. Si se encuentra alguna coincidencia, se comprueba que ese identificador con la misma dirección IP no esté dentro de *fakes.data*. En el mejor de los casos, la nueva pareja (usuario, IP) es añadida al fichero *fakes.data*.
- A continuación, se lee la variable que contiene las direcciones IP que han sido almacenadas en *fake.data* y se cuenta el número de veces que se repite cada una.
- Finalmente, se escribe en *IPFakes.data* aquellas direcciones IP clasificadas como maliciosas. Siendo estas direcciones todas las almacenadas en *fakes.data* que al menos se repiten 3 veces o más para usuarios distintos.

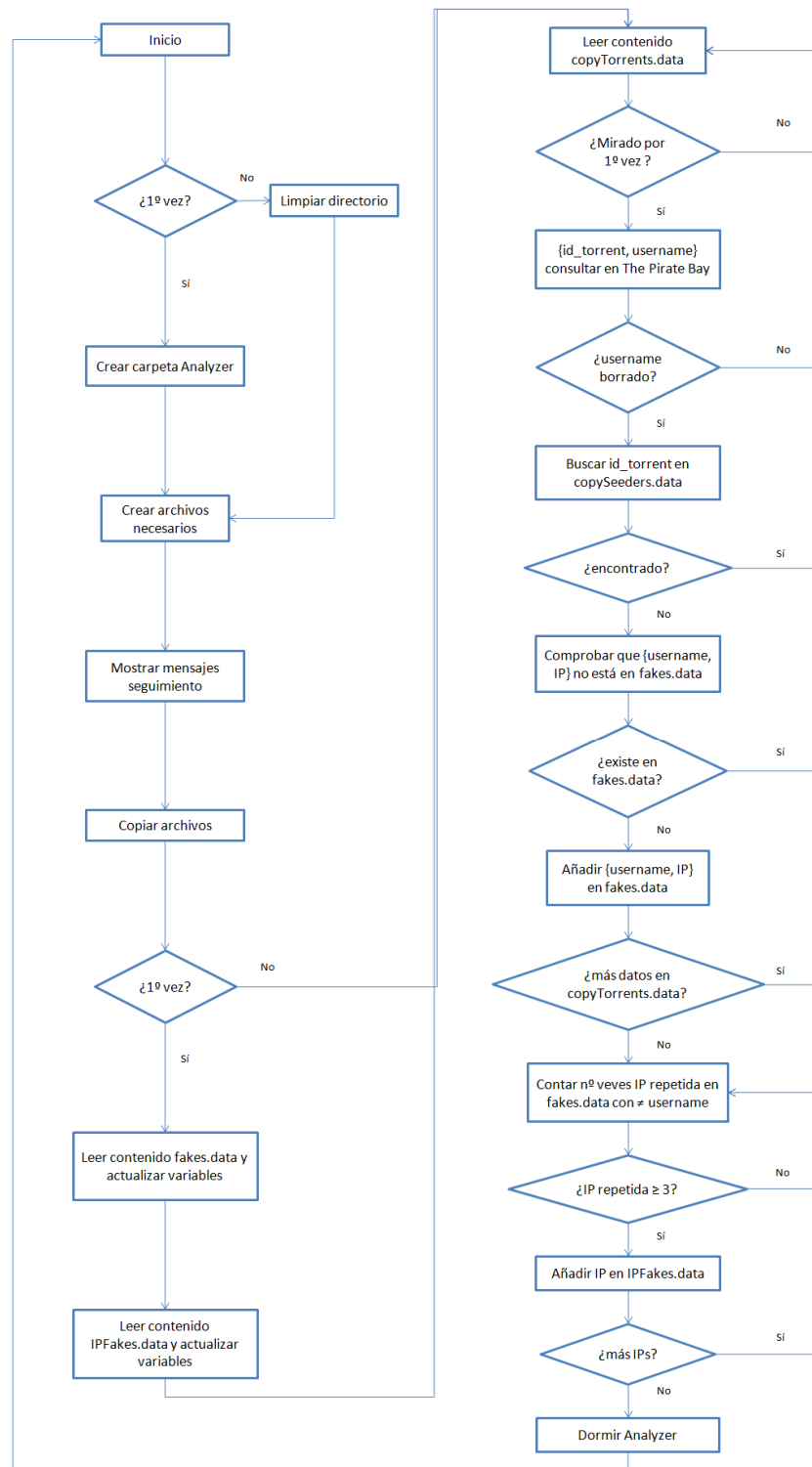


Figura 4. Diagrama de flujo de Analyzer.

Un usuario puede estar registrado en The Pirate Bay con un username y publicar torrents con una dirección IP concreta. El portal podría borrar a este usuario de la página y no por ello entender que este usuario es malicioso. El mismo usuario puede volver a registrarse y publicar torrents en la página, utilizando la misma dirección IP y un nombre de usuario distinto. Si este usuario vuelve a ser borrado de The Pirate Bay, la próxima vez que se registre con otro

username y vuelva a ser borrado, entonces sí, podrá considerarse que se está frente a un publicador malicioso.

La estrategia utilizada por este portal de BitTorrent para eliminar usuarios malintencionados, consiste en eliminar aquellas cuentas de usuarios de los que se ha recibido algún tipo de informe negativo como, por ejemplo: los comentarios de los propios usuarios del portal web, que indican que un username concreto está publicando contenidos malintencionados. Sin embargo, con este nuevo enfoque, Analyzer anticipa la detección de contenidos falsos.

Por tanto, gracias a la información recopilada por Crawler, se puede ir más allá y, trabajando con esta información almacenada, realizar un seguimiento para identificar aquellas direcciones IP maliciosas que suben contenidos corruptos a portales de BitTorrent como, por ejemplo, The Pirate Bay.

El problema de Analyzer es que la información que copia, así como la información que analiza y almacena, la gestiona a través de fichero de textos. Por lo tanto, las lecturas que hay que realizar sobre estos ficheros resultan ser lentas.

La solución a este problema pasa por usar bases de datos MySQL. Así, solucionar el problema de Analyzer y ser más eficiente. Los datos pasarán a estar mejor organizados y además se podrán manipular rápidamente pudiendo borrarlos y actualizarlos.

CAPÍTULO 4

TORRENT GUARD

4.1 INTRODUCCIÓN

Este capítulo resume la herramienta desarrollada previamente en otro proyecto, con la que se puede realizar la detección precoz de contenido falso en BitTorrent: Torrent Guard. Para entender un poco mejor por qué surge la necesidad de una herramienta así, se plantea el escenario actual que ha motivado el desarrollo de esta aplicación. El siguiente capítulo está basado en el artículo [20].

4.2 SITUACIÓN ACTUAL

Como se desprende de algunos estudios realizados [20], BitTorrent se ve sometido a una constante infección. Según estos estudios hasta el 30% de los contenidos publicados están asociados a contenidos falsos. Los usuarios de BitTorrent están expuestos a sufrir algún tipo de estafa o infección al descargar estos contenidos.

Por este motivo surge el desarrollo de Torrent Guard. Una herramienta de detección precoz de contenido falso en BitTorrent. Mientras que la mayoría de los portales identifican a los usuarios a través de su username, Torrent Guard es capaz de analizar la dirección IP de los publicadores y clasificarlos como malintencionados o no. Supone un cambio de enfoque a la

hora de detectar aquellos contenidos falsos o dañinos, y a los propios usuarios que realizan esta actividad.

Por tanto, analizar el comportamiento de estos usuarios y caracterizarlos es fundamental para entender cómo se producen estas publicaciones fraudulentas. Detectar y estudiar aquellos parámetros que identifican biunívocamente los usuarios malintencionados es un objetivo esencial.

4.3 METODOLOGÍA DE MEDIDA

Es necesario realizar un seguimiento de los archivos *.torrent* publicados en los portales web. Estos archivos almacenan toda la información acerca del contenido publicado. Al mismo tiempo, es imprescindible conocer los nuevos *.torrent* que se vayan publicando para estar al tanto de las novedades que se producen. Por tanto, se necesita identificar y caracterizar las principales propiedades de los publicadores. Estas actividades son llevadas a cabo por la aplicación Crawler, explicada en el Capítulo 3.

Dentro del amplio abanico de portales web dedicados a la publicación de archivos *.torrent*, Crawler rastrea The Pirate Bay. Es el portal web más utilizado por los publicadores de contenido debido a su notable popularidad, ofreciendo a los usuarios gran variedad de archivos para compartir.

Por tanto, una vez se obtienen las direcciones IP y los nombres de usuario asociados a los *.torrent* publicados, Crawler se conecta al tracker indicado. En este punto se pueden plantear distintos escenarios:

- El tracker sólo informa con la dirección IP del seeder inicial. Es posible si la conexión al enjambre se produce poco después de la publicación del *.torrent* en el portal web.
- El tracker anuncia la presencia de un seeder y pocos leechers del enjambre. Después, realizando conexiones a cada uno de los peers, se analiza el mensaje bitfields intercambiado en el enjambre.
- En el resto de casos, el tracker anuncia la existencia de varios seeders y otros tantos leechers en el enjambre, lo que hace inviable la identificación del seeder inicial.

A continuación es necesario determinar si los seeders identificados son malintencionados o no. Para ello, se ha desarrollado un segundo módulo que se conecta periódicamente a The Pirate Bay. Este nuevo módulo consulta al portal si algunos de los seeders identificados han sido borrados. En caso afirmativo, se tiene en cuenta que dicho seeder ha sido identificado como malintencionado y se guarda su dirección IP.

4.4 CARACTERIZACIÓN DE LOS PUBLICADORES DE CONTENIDO FALSO

BitTorrent es una de las aplicaciones que más tráfico de datos genera en la actualidad. Su mayor dificultad es poder frenar o mitigar la constante infección que sufren sus archivos. Para poder conseguir este objetivo es necesario identificar y caracterizar el comportamiento de sus publicadores. En concreto, centrar la atención en aquellos que comparten contenidos malware intencionadamente.

4.4.1 NÚMERO Y CONTRIBUCIÓN DE LOS PUBLICADORES MALINTENCIONADOS

En la actualidad son muchos los usuarios registrados en The Pirate Bay que comparten recursos a través del portal. En el estudio realizado en [20] se ha detectado que son pocas las direcciones IP responsables de la publicación de gran número de contenidos falsos. En media, unas 71 direcciones IP serían responsables de la publicación 4779 torrents malintencionados dentro del periodo de estudio de 14 días.

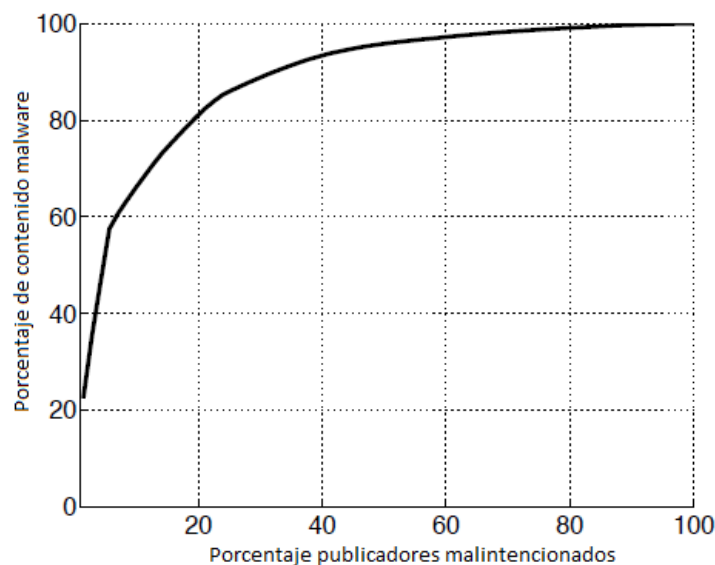


Figura 5. Porcentaje de contenido malware en función del porcentaje de publicadores malintencionados [20].

La Figura 5 representa el porcentaje de publicaciones falsas, en función del porcentaje de publicadores malintencionados. Como se observa, para pequeños porcentajes de publicadores malintencionados se obtienen grandes porcentajes de contenido falso. Por ejemplo, el 14% de los publicadores malintencionados (10 direcciones IP), serían responsables del 75% de los contenidos fraudulentos.

Al aumentar gradualmente el porcentaje de publicadores, se observa un incremento en el porcentaje de contenido falso. Pero el porcentaje de estos contenidos falsos sigue siendo especialmente alto para porcentajes bajos de publicadores. En concreto, un 20% de los

publicadores malintencionados o, lo que sería lo mismo, 14 direcciones IP son responsables de la publicación de más del 80% de los contenidos falsos.

Por tanto, se puede decir que un número reducido de usuarios registrados en The Pirate Bay son los autores de la constante intoxicación que está sufriendo el entorno de BitTorrent.

4.4.2 UBICACIÓN DE LOS PUBLICADORES MALINTENCIONADOS

Realizando una consulta a través del portal web desarrollado (ver Capítulo 5), para consultar cuáles son los principales proveedores de servicio de Internet, se obtiene lo siguiente:

Position	ISP	Country	Number of torrents
1	Comcast Cable	US	748
2	Ovh Systems	FR	619
3	Road Runner	US	481
4	Verizon Internet Services	US	263
5	Charter Communications	US	207
6	TeliaSonera AB	SE	205
7	NIB (National Internet Backbone)	IN	199
8	Altibox AS	NO	187
9	Com Hem AB	SE	180
10	Telefonica de Espana	ES	174

Figura 6. Principales ISP utilizados por usuarios de BitTorrent.

De los 10 ISP mostrados en la Figura 6 encabezan la lista: Comcast Cable [43], Ovh Systems [25] [26] y Road Runner [44]. Estas empresas se dedican principalmente a ofrecer servicios de *hosting*⁽¹⁰⁾. Es decir, a través de estas empresas se pueden alquilar servidores dedicados que permite alojar un espacio propio en Internet. Se aprecia que el número de *.torrent* publicados a través de estos ISP supera ampliamente el número de archivos publicados utilizando otros ISP.

En el estudio realizado en [20], se mapearon las direcciones IP de cada uno de los principales publicadores de contenido con el ISP correspondiente. Se obtuvo que 17 de los 20 principales publicadores que se examinaron, operaban desde proveedores de hosting.

De este comportamiento se deduce que los publicadores de contenidos falsos buscan: por un lado, recursos que les permitan llevar a cabo su actividad y distribuir el contenido de manera segura y cómoda; y, por otro lado, garantizar su anonimato.

⁽¹⁰⁾ Hosting: alojamiento web.

4.4.3 UTILIZACIÓN DE LAS CUENTAS EN THE PIRATE BAY

Con el fin de prohibir la generación automática de cuentas de usuario, en The Pirate Bay se solicita al usuario que, a lo largo del proceso de creación de una cuenta, introduzca un código CAPTCHA⁽¹¹⁾. Por lo tanto, los usuarios que quieran registrarse, entre ellos los publicadores malintencionados, están obligados a crear las cuentas de forma manual.

Como se expone en el estudio [20], el número de cuentas generadas en media por los publicadores malintencionados es de 4 al día. A través de las cuales se estarían inyectando contenidos falsos. Otra estrategia menos popular, mencionada en el mismo estudio, es que algunos publicadores malintencionados utilizan las cuentas de otros usuarios que no están bajo sospecha. Así, gestionar su actividad desde estas cuentas y estar más tiempo publicando contenidos falsos.

4.4.4 ESTRATEGIAS DE PUBLICACIÓN

Son dos las principales estrategias utilizadas por los publicadores de contenidos falsos dentro de The Pirate Bay:

- Están aquellos que inmediatamente después de registrarse dentro del portal, publican grandes cantidades de contenido falso. Cuando son detectados y su cuenta se elimina del portal, crean otra y repiten el proceso.
El uso de esta técnica es secundado por el 70% de los publicadores.
- La otra estrategia contempla aquellos usuarios más conservadores. Son aquellos que publican pocos archivos falsos, uno o dos, utilizando el mismo nombre de usuario. Así retrasan su detección y eliminación del portal de BitTorrent.
El 30% de los usuarios se ven atraídos por esta estrategia.

Es notable la diferencia de tiempo que necesita el portal web para detectar a los usuarios que utilizan una técnica u otra. Aquellos usuarios que utilizan la primera estrategia de publicación son detectados y borrados en un promedio de 92 minutos. Mientras, aquellos que utilizan la segunda son eliminados en 253 minutos.

⁽¹¹⁾ Código CAPTCHA: secuencias de números y/o letras que se utilizan para evitar que una máquina pueda gestionar de forma automática, por ejemplo, el registro en una página web.

4.4.5 ESTRATEGIAS PARA ATRAER A DOWNLOADERS

Los publicadores malintencionados tienen como objetivo conseguir que el mayor número de descargas se produzca sobre aquellos elementos que ellos publican. Así, con el objetivo de captar la atención de los usuarios, llevan a cabo distintas tácticas:

- Asignar a los *.torrent* nombres populares. Por ejemplo: el título de alguna película estrenada recientemente.
- Simular que el contenido ha sido publicado por un usuario de confianza. Por ejemplo: añadir la firma de un usuario de confianza en el título del archivo que quieren filtrar como malicioso.
- Presentar estadísticas atractivas del rendimiento del *.torrent* falso. Por ejemplo: elevar el número de seeders o leechers que poseen el archivo.

Los publicadores malintencionados son conscientes del funcionamiento del entorno de BitTorrent. Por eso, para poder llevar a cabo su actividad y poder mantener su anonimato, publican el contenido falso a través de servidores de hosting. El uso de estos servicios y otras estrategias sofisticadas, les garantiza poder realizar su actividad de forma exitosa.

4.5 PERFILES DE LOS PUBLICADORES DE CONTENIDO MALINTENCIONADO

El análisis de los datos recopilados en el estudio [20], revela la presencia de 3 perfiles diferentes entre los publicadores de contenido. Por un lado, están aquellos que propagan contenido malintencionado. Por otro lado, aquellos que aprovechan esta actividad para estafar a los usuarios. Finalmente, las agencias antipiratería.

A continuación se describen cada uno de estos perfiles:

- Propagadores de contenido malintencionado.
Son usuarios que tiene la intención de propagar archivos falsos por sí mismos, entre los usuarios de BitTorrent. Estos contenidos están manipulados intencionadamente, pretendiendo hacerlos pasar por otros sanos. Estos archivos se ejecutan de manera errónea. Es decir, una vez son descargados no funcionan de forma correcta, pidiendo al usuario que realice cambios en su equipo o no almacenan la información que se quería descargar. Estos archivos son potencialmente dañinos.
- Estafadores.
El objetivo de este grupo de usuarios va más allá que el descrito anteriormente. Partiendo de la misma iniciativa que los anteriores, los estafadores pretenden infectar a los usuarios y además beneficiarse de ello.

Por ejemplo, un comportamiento habitual de este grupo es: proporcionar archivos protegidos por contraseñas. Al mismo tiempo, facilitan enlaces a alguna página web a través de la cual obtener la clave. En este punto, llegan a solicitar al usuario: alguna tarjeta de crédito para poder completar la visualización de la descarga o, completar algún formulario con sus datos personales que conlleva la suscripción encubierta en algún servicio de mensajería corta que se cobra al usuario.

▪ Agencias antipiratería.

Las agencias antipiratería publican versiones falsas de los contenidos con derecho de autor que quieren proteger. Al contrario que los grupos anteriores, no tienen un propósito fraudulento.

Las copias introducidas por las agencias antipiratería son limitadas e introducidas intencionadamente en periodos concretos. Por ejemplo: semanas antes y después de la publicación de una película o lanzamiento de un nuevo disco.

4.6 TORRENT GUARD

El enfoque de algunas investigaciones analiza la vulnerabilidad del protocolo BitTorrent. Otras, enfocan el problema desde el punto de vista de la falta de privacidad que ofrece el protocolo. Sin embargo, el planteamiento del presente Proyecto no es otro que abordar el problema que sufre el entorno de BitTorrent en relación con la infección de sus ficheros.

Se ha visto que las técnicas que utilizan los portales de BitTorrent para eliminar contenidos falsos o dañinos, son ineficientes y requieren la colaboración de las personas. Además, la eliminación de estos contenidos fraudulentos se lleva a cabo únicamente en aquellos portales donde se identifican. Por lo tanto, no se elimina por completo de todos los escenarios de BitTorrent.

Torrent Guard es una herramienta desarrollada con el objetivo de automatizar y acelerar el proceso de detección de publicadores malintencionados. Las principales ventajas de Torrent Guard son:

- Identifica al publicador malintencionado por su dirección IP.
- Permite identificar un contenido fraudulento justo después de su publicación.
- Es capaz de eliminar el contenido de todos los escenarios de BitTorrent: reportando el valor de *infohash*.
- Identifica el contenido fraudulento independientemente del portal del cual se ha descargado u obtenido el *.torrent*. Siempre que proceda del servicio DHT⁽¹²⁾ de BitTorrent.
- Puede prevenir la descarga de contenidos malware.

⁽¹²⁾ DHT: en inglés Distributed Hash Tables [28]. Las tablas de hash distribuidas son una clase de sistemas distribuidos descentralizados que proveen un servicio de búsqueda similar al de las tablas de hash, donde pares (clave, valor) son almacenados en el DHT, y cualquier nodo participante puede recuperar de forma eficiente el valor asociado con una clave dada.

- Su uso puede reducir el número de ordenadores infectados y estafas sufridas por los usuarios.

Por tanto, el particular enfoque de Torrent Guard no sólo permite detectar contenidos falsos de una forma precoz. El uso apropiado de esta aplicación también contribuye a la reducción del número de infecciones y estafas sufridos por los usuarios de BitTorrent.

4.6.1 IMPLEMENTACIÓN DE TORRENT GUARD

El funcionamiento de Torrent Guard es fruto de la unión de los distintos módulos de Crawler, que se han explicado anteriormente, y las nuevas pautas de análisis e identificación elaboradas.

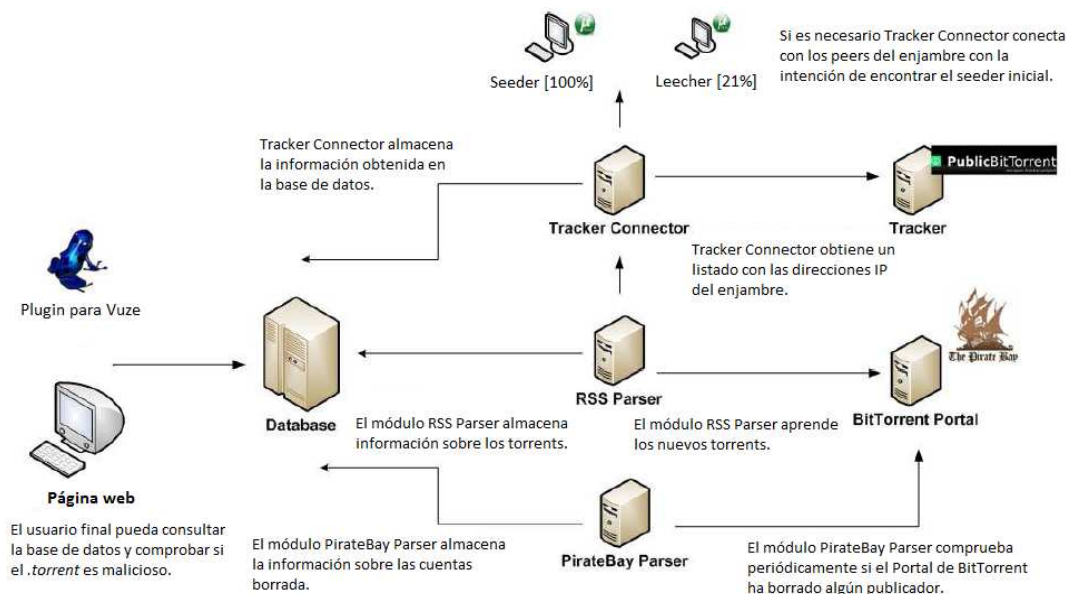


Figura 7. Esquema de Torrent Guard [20].

Torrent Guard utiliza The Pirate Bay con la intención de identificar nuevos publicadores corruptos y las direcciones IP desde las que operan. Con este fin, el analizador RSS continuamente monitoriza la distribución RSS de The Pirate Bay para aprender sobre los nuevos torrents e identificar el nombre de usuario para cada uno de ellos. Además, envía una pregunta al *Tracker Connector* que devuelve la dirección IP del seeder inicial (si es posible). Ambos, el nombre de usuario y la dirección IP, se almacenan en la base de datos. En paralelo, el analizador de The Pirate Bay conecta periódicamente con la página web de los diferentes publicadores descubiertos dentro de The Pirate Bay. Si se encuentra que la cuenta de un publicador ha sido borrada, este usuario y todos sus torrents son marcados como falsos. Además, se anota la dirección IP del publicador como una dirección potencialmente falsa. Si tres cuentas diferentes asociadas a la dirección IP son borradas de The Pirate Bay, se

considera esta dirección IP como dirección falsa. Desde ese momento, cualquier contenido publicado desde esta dirección IP es identificado justo en el momento de su publicación y reportado como falso.

Por lo tanto, en el peor de los casos, por ejemplo para un nuevo publicador malintencionado, Torrent Guard emplea el mismo tiempo que The Pirate Bay para identificar contenidos falsos. Sin embargo, una vez la dirección IP del publicador falso ha sido identificada, Torrent Guard es capaz de reportar contenidos falsos inmediatamente después de su publicación lo que proporciona una mejora significativa en comparación con los mecanismos de detección estándar. Además, con Torrent Guard no es necesario reportar cada contenido falso manualmente. Simplemente 3 informaciones podrían ser suficientes para identificar el usuario malintencionado y en consecuencia todos sus futuros torrents podrían ser automáticamente clasificados como malware.

Además, las soluciones que existen actualmente están limitadas al portal donde operan. Por ejemplo, en el caso de The Pirate Bay, una vez el contenido es identificado como falso es eliminado del portal web pero no de los escenarios de BitTorrent. Sin embargo, Torrent Guard es una solución válida para cualquier portal de BitTorrent, ya que es capaz de identificar la información hash del contenido falso evitando su descarga independientemente de la fuente desde donde el usuario obtuvo el archivo *.torrent*: cualquier portal BitTorrent o servicio DHT.

4.6.2 RENDIMIENTO DE TORRENT GUARD

Para medir el rendimiento de Torrent Guard se ha evaluado la capacidad de detección de la herramienta en comparación con el mecanismo de detección utilizado en The Pirate Bay. Los datos numéricos que se exponen a continuación han sido extraídos del estudio realizado en [20] durante los 14 días de prueba.

En primer lugar se han contado cuántos contenidos fraudulentos publicados en The Pirate Bay son identificados por Torrent Guard justo después de su publicación. Por otra parte, cuánto tiempo le lleva a The Pirate Bay identificar este contenido falso. Los resultados obtenidos muestran que Torrent Guard es capaz de detectar más pronto cerca del 50% del contenido falso publicado en The Pirate Bay.

Como se ha comentado anteriormente, la técnica de detección de The Pirate Bay solo afecta a las publicaciones realizadas en ese portal. Mientras que, la técnica empleada por Torrent Guard es una solución útil para todos los escenarios de BitTorrent. Por lo tanto, considerando todas las descargas producidas (asociadas con contenido malware), que se han llevado a cabo durante el periodo de 14 días de pruebas, Torrent Guard podría haber evitado 1.35 millones de descargas de contenido malware. Extendiendo este resultado para un año completo, se podría establecer que Torrent Guard sería capaz de eliminar más de 10 millones de descargas de contenido falso por año en comparación con la solución de The Pirate Bay y más de 35 millones si se consideran todos los escenarios de BitTorrent. Esto significa, dependiendo del éxito de las estrategias de los publicadores de contenido falso, la prevención de cientos de miles de infecciones malware e incidentes de estafa por año.

Por lo tanto, la evaluación inicial llevada a cabo en esta investigación muestra resultados muy prometedores para incentivar a la comunidad de BitTorrent a usar Torrent Guard.

4.6.3 LA EFICIENCIA DE TORRENT GUARD

Un sistema de detección, como es Torrent Guard, se caracteriza típicamente por el número de veces que se obtienen falsos negativos y falsos positivos por parte del sistema. Los falsos negativos están representados por aquellos torrents falsos que no son detectados. Los falsos positivos hacen referencia a aquellos contenidos clasificados como falsos y que no lo son.

Como se indica en [20], no es posible medir adecuadamente dicho parámetro, ya que requeriría una inspección manual de una gran cantidad (miles) de contenidos clasificados como legales (no falsos).

En su lugar, se realiza una evaluación abordable mediante la descarga de algunas docenas de torrents clasificados como legales por Torrent Guard. El análisis de los mismo confirma que la herramienta los ha clasificado correctamente, es decir, son verdaderos positivos. Sin embargo, de este resultado no se puede concluir que Torrent Guard sea infalible y detecte todo el contenido falso publicado en The Pirate Bay.

Con el fin de evaluar la tasa de falsos positivos de Torrent Guard, se analizan aquellos usernames no eliminados de The Pirate Bay, pero cuyo contenido ha sido clasificado por Torrent Guard como falso. Para ello, se descarga el contenido de cada uno de estos usuarios. Como consecuencia se obtiene que estos contenidos ciertamente están bien clasificados por Torrent Guard y que es el portal el que todavía no los ha detectado.

Por lo tanto, esta evaluación en pequeña escala demuestra que Torrent Guard tiene una tasa baja tanto en falsos positivos, como en falsos negativos.

CAPÍTULO 5

PÁGINAS WEB

5.1 INTRODUCCIÓN

En este capítulo se presenta la estructura de las páginas web que constituyen el portal. Se describe qué información se puede encontrar en cada una de las secciones. Inicialmente se explica cuál fue el punto de partida antes de desarrollar la estructura definitiva del portal web.

Al final del capítulo se mencionan algunos conceptos básicos a la hora de desarrollar una página web.

5.2 PUNTO DE PARTIDA: FAKE DETECTOR

Inicialmente se contaba con una estructura básica a partir de la cual se comienza a trabajar. Esta página web recibía el nombre de *Fake Detector*. La estructura de la página era bastante sencilla, como se muestra en la Figura 8. El comportamiento de la misma era básico y funcional, permitiendo al usuario consultar el estado de un archivo *.torrent* antes de comenzar con la descarga del mismo.

La página principal estaba desarrollada utilizando HTML. De esta forma, se definían los elementos básicos para que la página tuviese una buena estructura. Por otra parte, también se definía un archivo *.php* a través del cual se gestiona la consulta a la base de datos y así poder mostrar el resultado.

Sin más particularidades un usuario seleccionaba el archivo *.torrent* que quería analizar y a continuación se mostraba el resultado de la consulta de una forma clara y directa. El resultado podía ser: positivo, lo que quería decir que el archivo *.torrent* se correspondía con el contenido que se iba a descargar; negativo, indicando al usuario que el contenido era falso; o incierto, lo que implicaba que el archivo *.torrent* podía o no ser correcto.



Figura 8. Página de inicio para Fake Detector.

El objetivo de esta página era claro: establecer una forma práctica y sencilla de consultar si un archivo *.torrent* podía estar infectado o no.

5.3 ESTRUCTURA DEL PORTAL WEB

El objetivo de esta parte del Proyecto ha sido: desarrollar, evolucionar y mejorar Fake Detector. Creando un portal web funcional, sencillo y útil, que englobase las herramientas desarrolladas y unificase el acceso a MYPROBE y Torrent Guard. Que además mostrase información de interés relacionada con BitTorrent como, por ejemplo, artículos y otras publicaciones elaborados, así como avances dentro de las líneas de investigación.

El portal web está dividido en diferentes secciones o apartados. Cada uno de ellos cuenta con un enlace propio tanto en el menú principal como en la parte inferior del portal. Dentro de algunos apartados se han incluido subapartados con el fin de ofrecer un espacio individual en función del criterio de búsqueda elegido por el usuario.

Todas las secciones mantienen el mismo estilo gracias al menú principal que aparece en todas ellas. Así, los usuarios tienen la sensación de continuidad mientras navegan por el portal web y resulta bastante intuitivo encontrar información tanto de los publicadores de contenido, como sobre los archivos *.torrent*. Además, los enlaces a los apartados principales del portal se repiten al final del mismo, para que el usuario tenga distintas formas de acceder a ellos.

Dentro de cada apartado y subapartado, se explica al usuario qué información puede consultar. Además, se indica cómo debe rellenar los formularios para obtener la información. En cualquier caso, si surge algún problema durante una consulta, las páginas web muestran

mensajes de ayuda, con el fin de guiar al usuario para que pueda completar la consulta correctamente.

De forma esquemática, la estructura principal del portal web se muestra en la siguiente figura donde los rectángulos con fondo azul representan los apartados principales mostrados en el menú, que están siempre visibles en las distintas páginas. Mientras, los rectángulos con fondo blanco, representan los distintos subapartados.

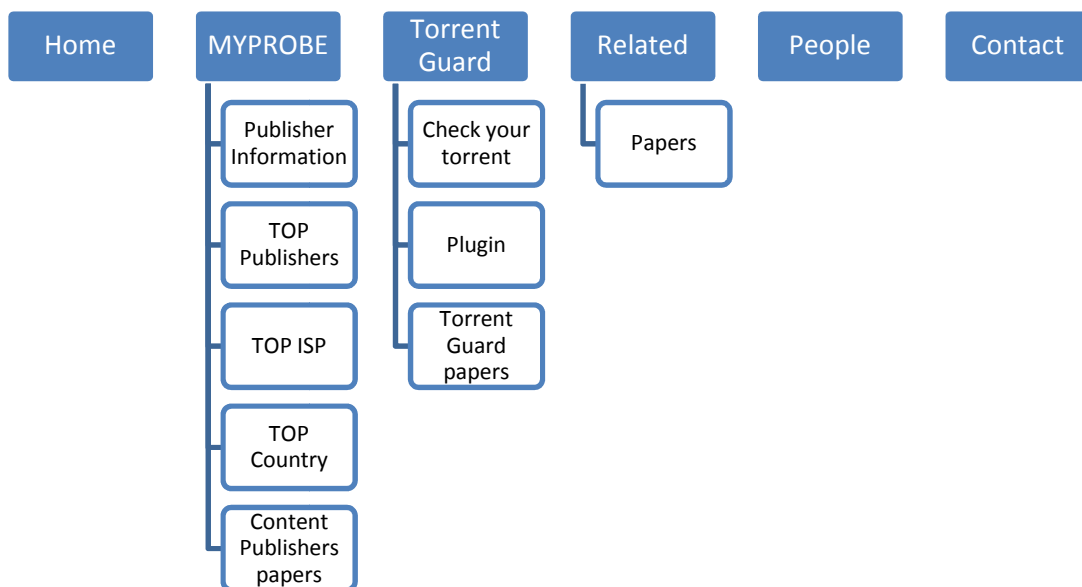


Figura 9. Estructura principal del portal web diseñado.

A continuación se describe la funcionalidad de cada parte y la información disponible en cada una de ellas.

5.3.1 HOME

Para acceder al portal web es necesario escribir en la barra de direcciones lo siguiente: <http://bittorrent.netcom.it.uc3m.es/>. Esto dirigirá al usuario a la página principal.

En este apartado se presenta el portal web con una breve introducción sobre el Proyecto en el que se ha estado trabajando. Como se indicaba anteriormente, en el menú principal se aprecian los distintos apartados, así un usuario puede visualizar rápidamente las opciones que tiene la página. Además, los enlaces principales se repiten en la parte inferior con el fin de facilitar a los usuarios la navegación a través del portal web.

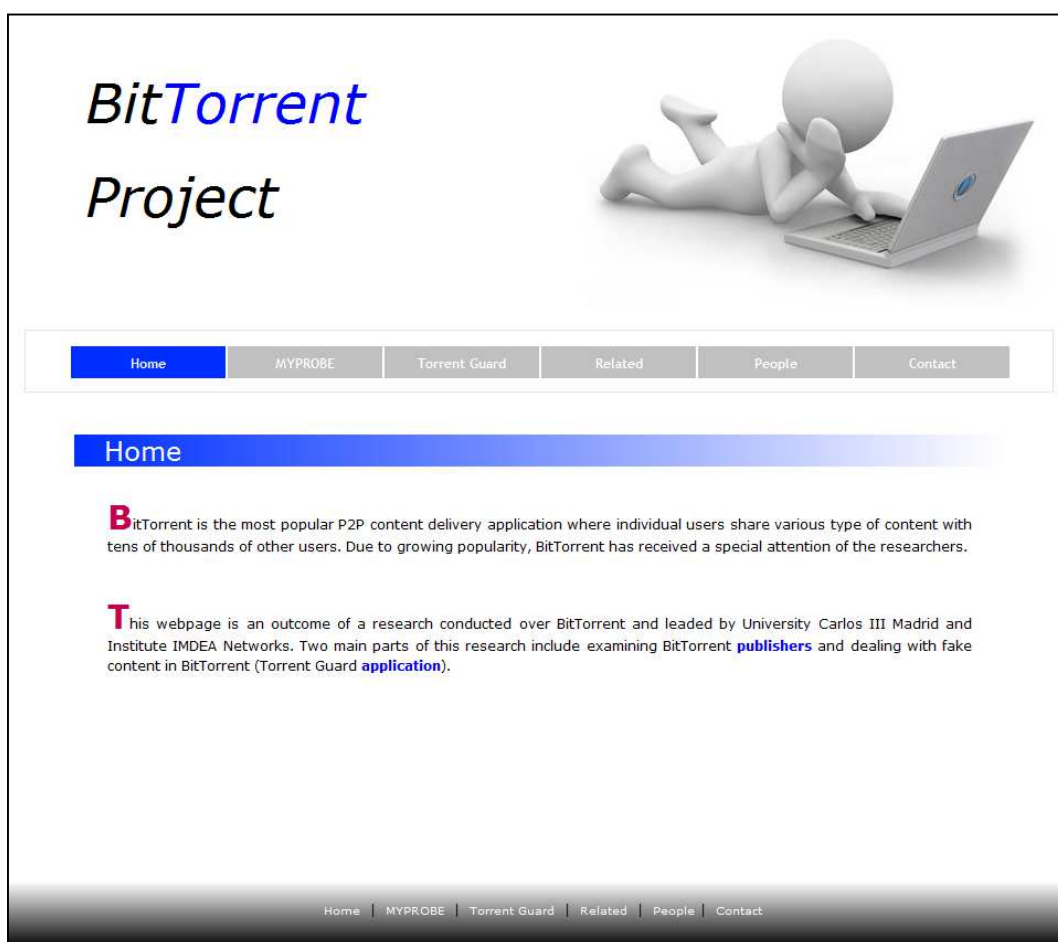


Figura 10. Página principal.

5.3.2 MYPROBE: MONITORING, IDENTIFYING AND PROFILING BITTORRENT PUBLISHERS

En este apartado un usuario puede conocer información relevante acerca de los publicadores de contenido de uno de los portales más importantes dentro de BitTorrent como es The Pirate Bay. Por un lado, el enlace principal nos introduce y presenta las opciones básicas que se pueden consultar en este apartado. Por otro lado, se ofrecen al usuario distintos subapartados que focalizan el criterio de búsqueda. De esta manera se plantean diferentes opciones para encontrar información acerca de aquellos publicadores de contenido:

- **Publisher Information:** Aquí se ofrece la opción de introducir un nombre de usuario cualquiera. Si no se introduce ningún nombre y se envía la petición, la página web muestra un mensaje al usuario indicándole que debe introducir el dato correctamente. Si se introduce un nombre de usuario pueden suceder dos cosas: que la consulta obtenga algún resultado o que por el contrario, no exista ninguna coincidencia con el criterio indicado. En el primer caso, se muestra la información oportuna y, en el segundo caso, se indica al usuario que no se ha encontrado ninguna información. El resultado de la consulta muestra al usuario el número de torrents publicados por el usuario en cuestión, así como las distintas categorías en las que se clasifican estos contenidos. Igualmente, se muestran las direcciones IP de los seeders principales. Es decir, aquellas direcciones IP asociadas con el username especificado. En algunas ocasiones puede darse el caso de identificar el número de torrents y categorías en las que publica un usuario. Sin embargo, no poder averiguar las direcciones IP a través de las cuales publica contenidos. En cualquier caso, la página de resultados se adapta en función del escenario de respuesta.
- **TOP Publishers:** El usuario puede consultar los principales publicadores de contenido que aparecen dentro de The Pirate Bay. Para ello se debe indicar al menos el número de publicadores principales que se quiere ver, pudiendo filtrar el resultado por una categoría concreta dentro de las que se ofrecen en el panel desplegable. Igual que sucede en el apartado anterior, se muestran mensajes al usuario para ayudarle en todo momento a completar la consulta de forma satisfactoria. El resultado muestra una tabla con la información solicitada. En función del número indicado el resultado puede aparecer en diferentes páginas, es decir, paginado. No obstante, siempre existe la opción de mostrar una única tabla con todas las entradas disponibles. En cada línea de la tabla se muestra el nombre del usuario que publica contenido en The Pirate Bay, el número de torrents publicados por el mismo y un enlace con más información. El resultado está organizado por número de torrents publicados siendo el primero el que más torrents ha publicado. Si se selecciona el enlace *More information*, la página muestra la información para ese usuario disponible en Publishers Information. Es decir, aparece la misma información que se hubiera obtenido introduciendo el nombre de usuario en el apartado anterior: Publishers Information. Si se selecciona el enlace ubicado en el nombre del usuario, se abre una nueva ventana que redirige al usuario a la página de dicho usuario en The Pirate Bay.

- **TOP ISP:** Este subapartado dentro de MYPROBE sigue la misma línea que TOP Publishers. Muestra el resultado en una tabla organizando la información en función del número de torrents de mayor a menor. Igualmente cabe destacar que el resultado puede aparecer paginado o en una única tabla, esto dependerá del criterio de búsqueda.

Para iniciar la consulta, es necesario indicar como mínimo el número de ISP principales que se quieren conocer, estando fijado el límite superior en 100. También se da la opción al usuario de filtrar la consulta en función de una categoría concreta (a seleccionar de la lista desplegable).

La primera columna muestra la posición que ocupa un ISP concreto dentro de los parámetros de búsqueda elegidos. La segunda columna de la tabla de resultados muestra el nombre del proveedor de servicio de Internet que se ha utilizado para publicar los torrents. A continuación, la tercera columna indica el país asociado a ese ISP y finalmente, en la cuarta columna aparece el número de torrents que han sido publicados a través del ISP.

- **TOP Country:** Igual que en los dos subapartados anteriores, TOP Country permite al usuario listar los principales países que tienen un elevado tráfico en lo que a publicación de torrents se refiere. Pudiendo filtrar la consulta por categorías, indicando simplemente el número de países principales que se quieren conocer, el resultado se muestra en una tabla. Al lado de cada país se indica el número de torrents identificados para el mismo.
- **Content Publishers papers:** Este espacio ha sido reservado para poder publicar aquellos artículos que tengan una relación directa con los publicadores de contenido en los portales de BitTorrent. El objetivo es publicar aquellos documentos que permitan a los usuarios del portal web estar al tanto de las últimas novedades.

En el Capítulo 6 se muestran ejemplos de consultas realizadas sobre la página web.

5.3.3 TORRENT GUARD

Este apartado engloba la herramienta desarrollada para la detección de contenido malintencionado publicado dentro de BitTorrent, explicada en el Capítulo 4. Un usuario encontrará las herramientas necesarias de consulta para poder identificar si un contenido concreto puede ser malintencionado.

Así como el apartado MYPROBE está más enfocado a consultar información acerca de los publicadores, este apartado está enfocado a consultar el estado de aquellos *.torrent* que se han descargado. De ahí que Torrent Guard esté presente como una herramienta de consulta web para los usuarios, como también en los distintos plugin desarrollados para distintos clientes BitTorrent. En concreto estos plugin han sido desarrollados para Vuze y μ Torrent por Beatriz Calvo y Diego Cristina, respectivamente. Por tanto, todos los elementos quedan englobados dentro del mismo espacio web para facilitar al usuario el uso de los mismos.

En todo momento se explica al usuario la información mostrada a través de la página. De esta forma, en Torrent Guard se presentan las aplicaciones y en sus subapartados se explica a los usuarios qué tienen que hacer para poder comprobar la veracidad de los contenidos publicados en el entorno de BitTorrent. Mostrando los mensajes oportunos por pantalla en caso de fallo o imprecisión en los datos introducidos en la página web.

A continuación se describen brevemente los subapartados dentro de Torrent Guard:

- **Check your torrent:** Cuando un usuario descarga un archivo *.torrent* se le ofrece la posibilidad de comprobar el estado de este archivo. Así garantizar que el contenido está bien y no supone ningún riesgo.

En primer lugar, se indica el estado del *.torrent*. De esta forma se especifica si el archivo es: correcto, falso, incierto o si no se dispone de información suficiente para clasificarlo. Además, el resultado muestra al usuario información básica acerca del *.torrent* que está comprobando. Esta información es: el título, el identificador del *.torrent*, la fecha de creación, la categoría, subcategoría y el autor. Para este último se ofrece un enlace directo a la página personal que este publicador tiene en The Pirate Bay.

Por otro lado se ofrece otra información de interés relacionada con el *.torrent* que se está comprobando, como es: algún comentario acerca del archivo, el tamaño de los trozos que completan el *.torrent*, así como la información hash que identifica biunívocamente el archivo.

Finalmente, el resultado para aquellos torrents que han sido identificados como correctos muestra, si se dispone de la información, una tabla con los seeders principales identificados para el autor .

Por tanto, a través de esta consulta web, un usuario puede confiar en la veracidad o no del *.torrent* que ha descargado. Además, gracias a la información hash que identifica al archivo, se pueden consultar otros torrents que no hayan sido descargados del portal The Pirate Bay. Es decir, independientemente del portal de BitTorrent que se haya utilizado para descargar el archivo, la aplicación web es capaz de identificar y valorar el estado del *.torrent* que se esté consultando.

Debido a las evoluciones actuales relacionadas con los archivos *.torrent*, como son los enlaces magnéticos, se ha optado por incluir dentro de la página otros criterios de consulta de torrents. Además de poder verificar un archivo *.torrent* como se ha explicado anteriormente, se han añadido otras dos opciones. En primer lugar se ofrece al usuario la oportunidad de consultar el estado de un archivo *.torrent* simplemente introduciendo el valor de infohash del mismo. Es decir, incluso sin realizar la descarga del archivo, se puede informar a los usuarios sobre el estado del mismo y reportarles igualmente la información oportuna.

En segundo lugar, a través de la página se pueden introducir directamente enlaces magnéticos. Por lo tanto, la página está completamente adaptada a las necesidades de los usuarios, pudiéndose comprobar el estado de los torrents en distintos formatos.

- **Plugin:** Este subapartado dentro de Torrent Guard contiene las aplicaciones desarrolladas para distintos clientes de BitTorrent. Por un lado el plugin para Vuze desarrollado por Beatriz Calvo y por el otro, el plugin para μ Torrent desarrollado por Diego Cristina. A través de los enlaces un usuario del portal puede descargar las aplicaciones e instalarlas de forma sencilla y rápida en función del cliente de BitTorrent elegido.

De esta manera, no sólo se ofrece al usuario la posibilidad de consultar su archivo *.torrent* a través de la página web, sino que también se le facilitan aplicaciones desarrolladas con el mismo objetivo de detección de contenidos falsos.

- **Torrent Guard papers:** Igual que sucedía en MYPROBE, para Torrent Guard también se ha creado un espacio para compartir información de interés sobre la herramienta desarrollada. De esta forma se puede ofrecer al usuario información acerca de las líneas de investigación que se están siguiendo y los avances que surgen.

En el Capítulo 6 se muestran ejemplos de consultas realizadas sobre la página web.

5.3.4 RELATED

En apartados anteriores, MYPROBE y Torrent Guard, existen subapartados especiales dedicados a artículos e información de interés de ambas herramientas. Sin embargo, con el fin de dar a conocer otros trabajos publicados que siguen la misma línea de investigación, se ha reservado este espacio dentro del portal web para poder mostrar a los usuarios más información acerca de BitTorrent.

Los documentos publicados actualmente, por supuesto, abarcan investigaciones realizadas en el entorno de BitTorrent. Pero también ofrecen información acerca de los escenarios de estudio y de diferentes técnicas utilizadas para medir el tráfico en BitTorrent.

Las figuras siguientes muestran el aspecto que tiene este apartado dentro del portal web.



Figura 11. Apartado Related.

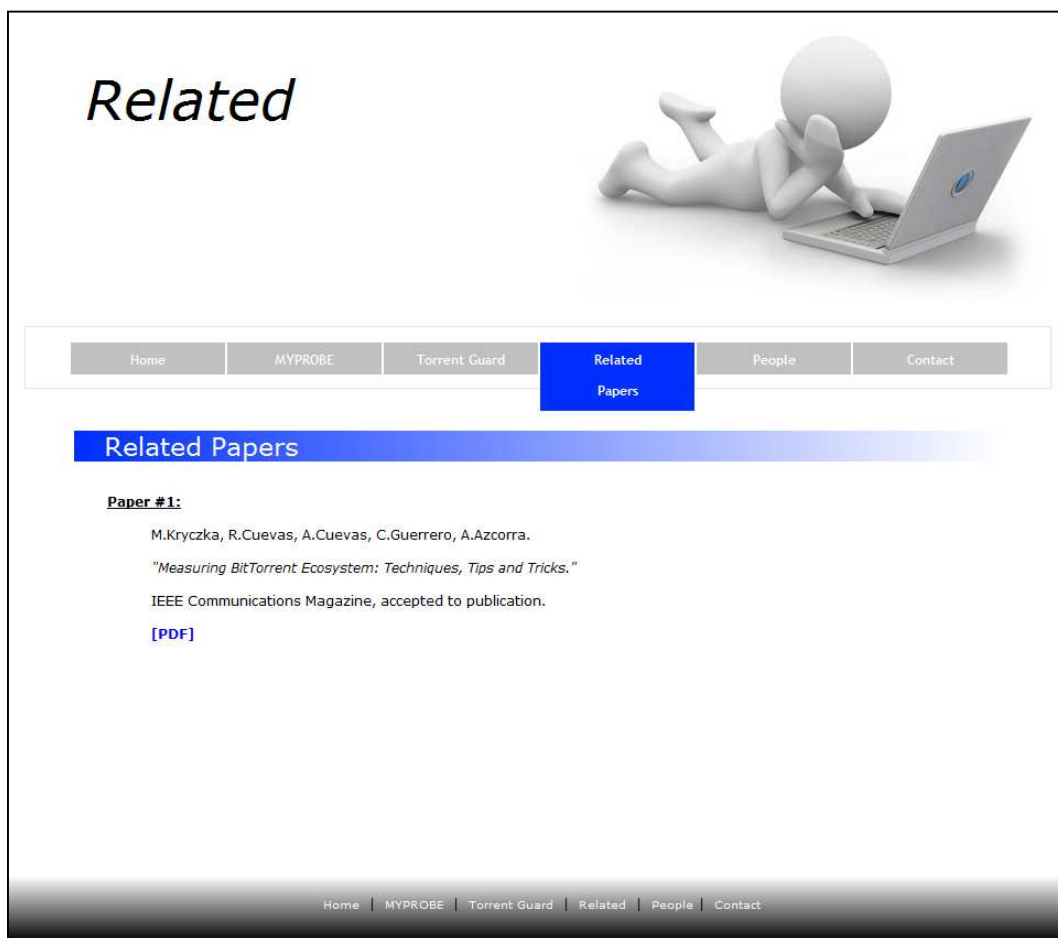


Figura 12. Subapartado Papers, enlaces disponibles.

5.3.5 PEOPLE

Este apartado es fundamental. En él se nombran a todas las personas que han participado de una forma u otra dentro del Proyecto. Desde las personas que han escrito los documentos y artículos que aparecen en la web, y han investigado acerca de BitTorrent. Hasta las personas que han desarrollado las aplicaciones mencionadas a lo largo de esta memoria.

Como reconocimiento al trabajo realizado, en este apartado se engloban a todos los que han participado en el desarrollo de este Proyecto y de otros ligados a él. De esta forma, queda constancia de la colaboración de los mismos dentro de este proceso.

La figura siguiente muestra el resultado que ofrece la página web:

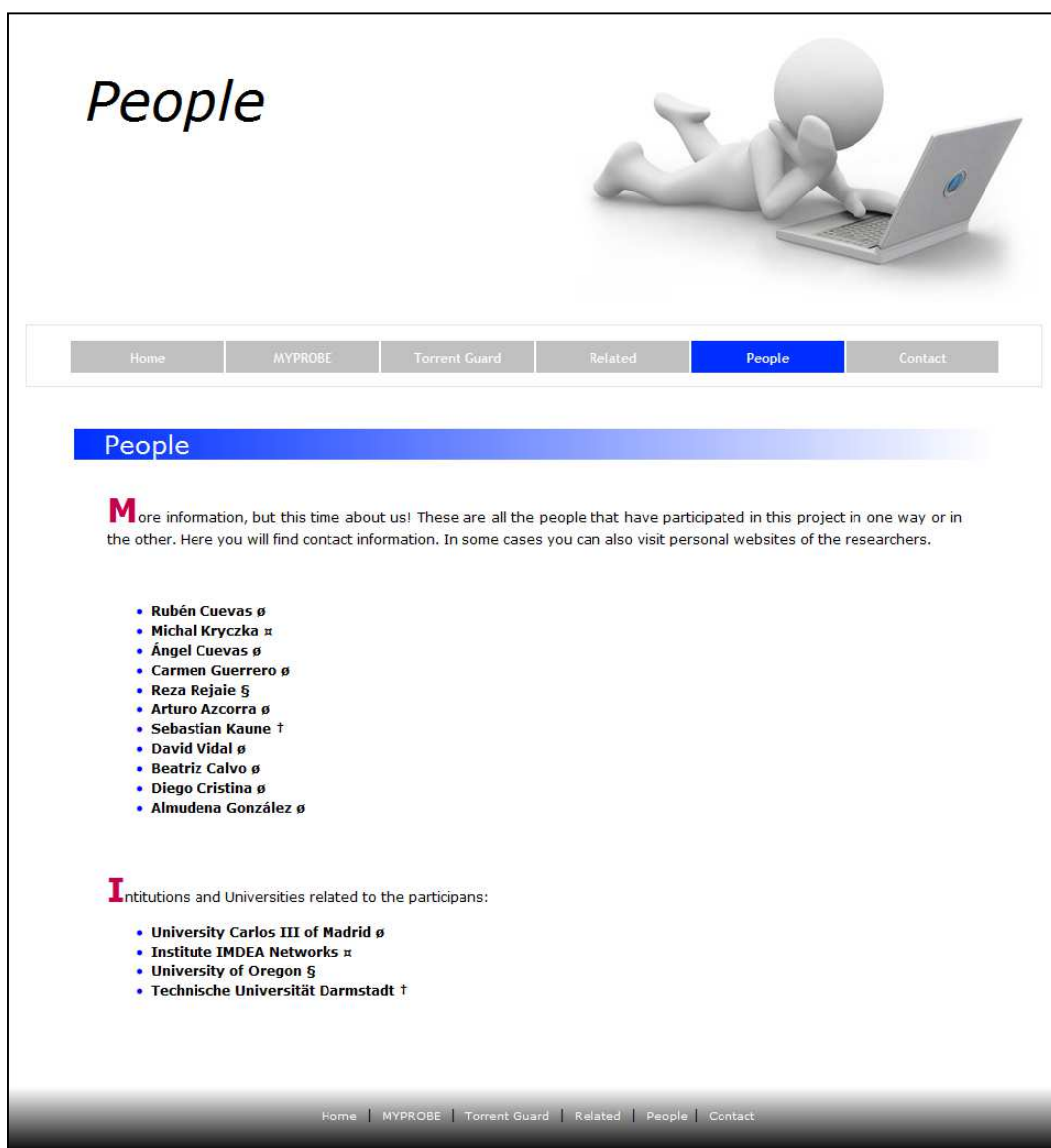


Figura 13. Apartado People.

5.3.6 CONTACT

El último apartado es: Contacto. Aquí se ofrece al usuario la posibilidad de escribir a los miembros del equipo y realizar otro tipo de consultas, formular dudas, exponer sugerencias o simplemente expresar su opinión.

Para ello el usuario debe rellenar un sencillo formulario web. A continuación, en la siguiente figura, se muestra el aspecto del mismo.

Contact

Home MYPROBE Torrent Guard Related People **Contact**

Contact

You can contact us through this place. Is quite easy, you just need to write your message and complete the e-mail box. This point is really important, we need you to check your e-mail before sending it. If it is wrong, we will not be able to answer.

Additionally you can also select a topic from the regarding dropdown menu and write your name. Both items are optional. Don't be shy, we are glad to answer your questions. We will answer as soon as posible.

REMEMBER: write a correct e-mail.

Name:

E-mail:

Regarding:

Message:

Home | MYPROBE | Torrent Guard | Related | People | Contact

Figura 14. Apartado Contact.

El usuario debe completar unos datos sencillos para formular su consulta. En primer lugar, se le pide que introduzca el nombre. A continuación, se solicita que introduzca una dirección de correo electrónico. En este caso, la página web comprueba la correcta escritura de la misma. Después, el usuario puede seleccionar un tema concreto dentro de las opciones ofrecidas en el panel desplegable. Finalmente, puede escribir el mensaje con la consulta, solicitud o petición que considere oportuna. En cualquier momento, si el usuario desea comenzar de nuevo, puede borrar el formulario seleccionando *Clear text*.

Cuando el usuario manda el formulario, la página web comprueba que la dirección de correo electrónico se ha escrito bien. En caso de existir algún error, muestra un mensaje por pantalla al usuario, solicitando que por favor introduzca una dirección correcta. La siguiente figura muestra el mensaje ofrecido al usuario.

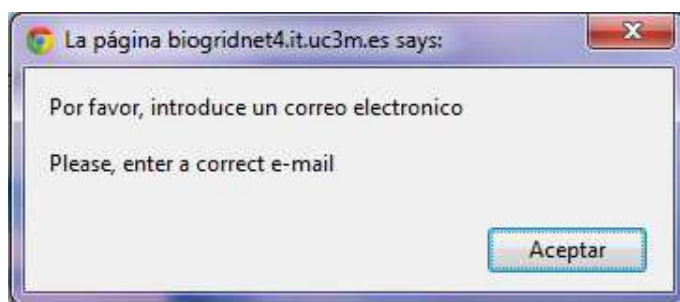


Figura 15. Mensaje de ayuda al usuario en el envío del formulario de contacto.

Si el usuario introduce una dirección de correo válida, se confirma el envío del mensaje mostrando la siguiente pantalla.

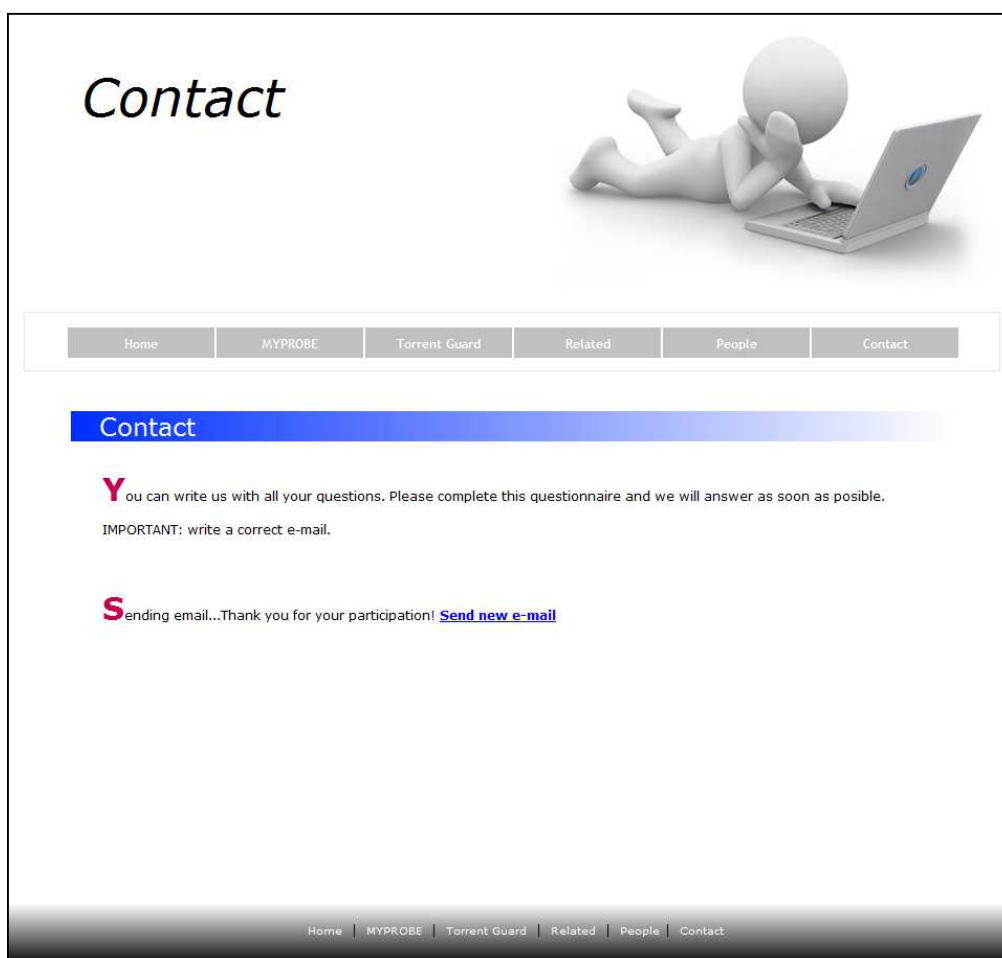


Figura 16. Confirmación de envío del mensaje.

5.4 HTML: HYPERTEXT MARKUP LANGUAGE

HTML [12], [13], [14], es un lenguaje de programación que permite escribir páginas web de forma estructurada. Al ser un lenguaje de hipertexto, es usado para describir la estructura y el contenido de una página web en forma de texto. Está compuesto por etiquetas que marcan el inicio y el fin de cada elemento del documento.

Los navegadores interpretan el código establecido en un fichero HTML y muestran al usuario la página web correspondiente. Además, estos archivos *.html* pueden contener referencias a otros elementos (audio, video, páginas web, *scripts*,...).

Algunos de los problemas que acompañan al lenguaje HTML es la diversidad de navegadores que hay en el mercado, los cuales no interpretan de la misma manera un código HTML. Esto obliga a comprobar uno a uno cómo se comportan los navegadores para un mismo código. Gracias al uso de hojas de estilo en cascada, Java Script u otras herramientas es posible organizar, optimizar, compatibilizar y automatizar el funcionamiento de las páginas web.

Para poder generar estos archivos basta con utilizar un editor de textos como, por ejemplo: Bloc de notas. Sin embargo, existen otros editores que cuentan con un entorno visual, de esta forma se puede ver en todo momento cómo va quedando la página web.

5.5 CONFLICTOS CON NAVEGADORES: HOJAS DE ESTILO EN CASCADA CSS

En la actualidad existen diferentes navegadores web a través de los cuales se puede consultar información en Internet. Algunos de los navegadores más populares son: Internet Explorer, Mozilla Firefox, Google Chrome, Opera o Safari.

A la hora de desarrollar una página web es muy importante saber que cada navegador interpreta de forma distinta las hojas de estilo en cascada que contienen las páginas web, como sucede con el código HTML.

Las hojas de estilo en cascada [10], [15],[16], son pautas que los navegadores deben interpretar para mostrar la estética de las páginas web de forma correcta. Las hojas de estilo garantizan mantener la homogeneidad en el estilo de las páginas web. Así, aquellas páginas que contengan una referencia a una hoja de estilo concreta, mostrarán la misma apariencia.

Por tanto, con las hojas de estilo en cascada se ayuda a que los navegadores interpreten correctamente el estilo de las páginas web y muestren el contenido de forma adecuada.

Uno de los objetivos de este Proyecto es que se pueda acceder al portal web desde todos los navegadores posibles. De esta forma, ofrecer la oportunidad de comprobar el contenido a un mayor número de usuarios. Por este motivo se han desarrollado varias hojas de estilo en función de las necesidades de los navegadores. Afortunadamente, las pautas descritas en las hojas de estilo son compartidas entre algunos navegadores. Por ejemplo, una hoja de estilo

para Mozilla Firefox es interpretada, a grandes rasgos, de la misma forma en Google Chrome. Sin embargo, el navegador Internet Explorer es la excepción y necesita una hoja de estilo individual.



Figura 17. Navegadores web.

El portal web desarrollado se visualiza perfectamente en: Google Chrome, Mozilla Firefox, Safari o Internet Explorer, siendo su visualización óptima si se usa Mozilla Firefox.

5.6 LENGUAJE DE PROGRAMACIÓN INTERPRETADO: PHP

PHP [17], [18], [19], es un acrónimo recursivo que significa *PHP HyperText Pre-processor*. Es un lenguaje de propósito general especialmente utilizado para el desarrollo de páginas web. Tiene gran aceptación debido a la simplicidad en su sintaxis, pudiendo realizar determinadas acciones de forma fácil y eficaz. Además posee capacidad para acceder a bases de datos de una manera sencilla.

Otro aspecto positivo de PHP es que puede ser embebido en páginas HTML. Es decir, se pueden escribir pequeños script en PHP a los cuales hacer referencia en las páginas HTML.

En el desarrollo del portal web se han desarrollado diferentes archivos *.php*. Cada uno especializado en una función distinta. Elaborando así códigos funcionales y modulares en función de la consulta a realizar. El objetivo principal de estos archivos ha sido establecer las consultas oportunas sobre las diferentes bases de datos a consultar.

CAPÍTULO 6

BÚSQUEDAS DENTRO DEL PORTAL WEB

6.1 INTRODUCCIÓN

En este capítulo se realizan distintas consultas en las páginas web con la intención de recorrer todos los escenarios de consulta/resultado existentes. De esta manera se pueden observar mejor cómo se muestran los resultados en cada uno de los distintos subapartados del portal, así como analizar los mismos.

Todas las consultas se realizan utilizando el navegador Mozilla Firefox. La versión utilizada es la 3.6.28, instalada en el ordenador dedicado a pruebas. El enlace de acceso directo a la página principal del portal web es: <http://bittorrent.netcom.it.uc3m.es>.

También existe la opción de acceder directamente a las secciones MYPROBE y Torrent Guard. Para ello basta con escribir en la barra de direcciones: <http://btpublishers.netcom.it.uc3m.es/> y <http://torrentguard.netcom.it.uc3m.es/>, respectivamente.

6.2 CONSULTAS REALIZADAS EN MYPROBE

En primer lugar se comenzará con un repaso de las consultas que se pueden realizar a través del apartado MYPROBE. De esta forma, ejecutando ejemplos sencillos sobre cada uno de los subapartados, repasar los resultados que se obtienen a través de las páginas web.

6.2.1 PUBLISHERS INFORMATION

Se inicia la consulta desde el apartado MYPROBE, accediendo al enlace Publisher Information que aparece en el menú principal. También se puede acceder a través del enlace *Details* dentro del texto del apartado MYPROBE.



Figura 18. Acceso a Publishers Information.

Se prueba qué sucede si no se introduce ningún nombre de usuario. Como se observa, en el resultado que aparece en la figura siguiente, la página detecta que no se ha introducido ningún usuario y muestra el siguiente mensaje de ayuda.



Figura 19. Mensaje de ayuda al usuario en Publishers Information.

Ahora se introduce el nombre de usuario sobre el que se va a realizar la consulta. En este caso se introduce un nombre inventado para conocer qué respuesta ofrece la página web.



Figura 20. Consulta de un usuario inventado en Publishers Information.

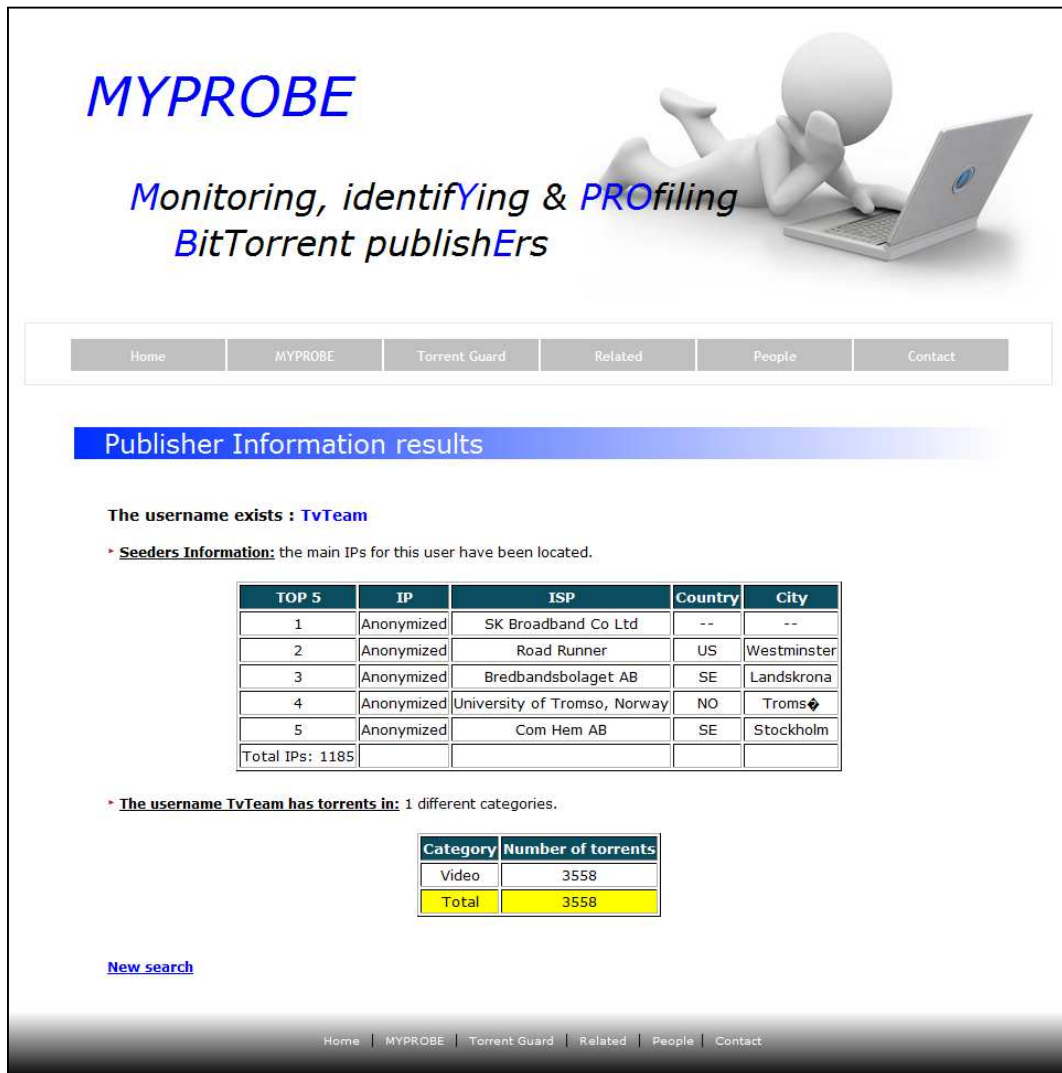


Figura 21. Resultado consulta de un usuario inventado en Publishers Information.

La página web detecta que el usuario introducido no produce ninguna coincidencia dentro de la base de datos en la que consulta. Por lo tanto, muestra al usuario el mensaje que aparece en la Figura 21.

Este resultado también aparece cuando no es posible mostrar información relevante del usuario sobre el que se está realizando la consulta. Esto puede suceder cuando el usuario ha sido borrado del portal de descargas The Pirate Bay y no ha sido posible realizar un rastreo completo del mismo.

A continuación se introduce un nombre de usuario identificado en The Pirate Bay: **TvTeam**. En el momento de realización de esta memoria **TvTeam** es el usuario principal entre los publicadores de contenido en The Pirate Bay.



MYPROBE
Monitoring, identifying & PROfiling
BitTorrent publishErs

Home MYPROBE Torrent Guard Related People Contact

Publisher Information results

The username exists : **TvTeam**

► **Seeders Information:** the main IPs for this user have been located.

TOP 5	IP	ISP	Country	City
1	Anonymized	SK Broadband Co Ltd	--	--
2	Anonymized	Road Runner	US	Westminster
3	Anonymized	Bredbandsbolaget AB	SE	Landskrona
4	Anonymized	University of Tromso, Norway	NO	Troms
5	Anonymized	Com Hem AB	SE	Stockholm
Total IPs: 1185				

► **The username TvTeam has torrents in:** 1 different categories.

Category	Number of torrents
Video	3558
Total	3558

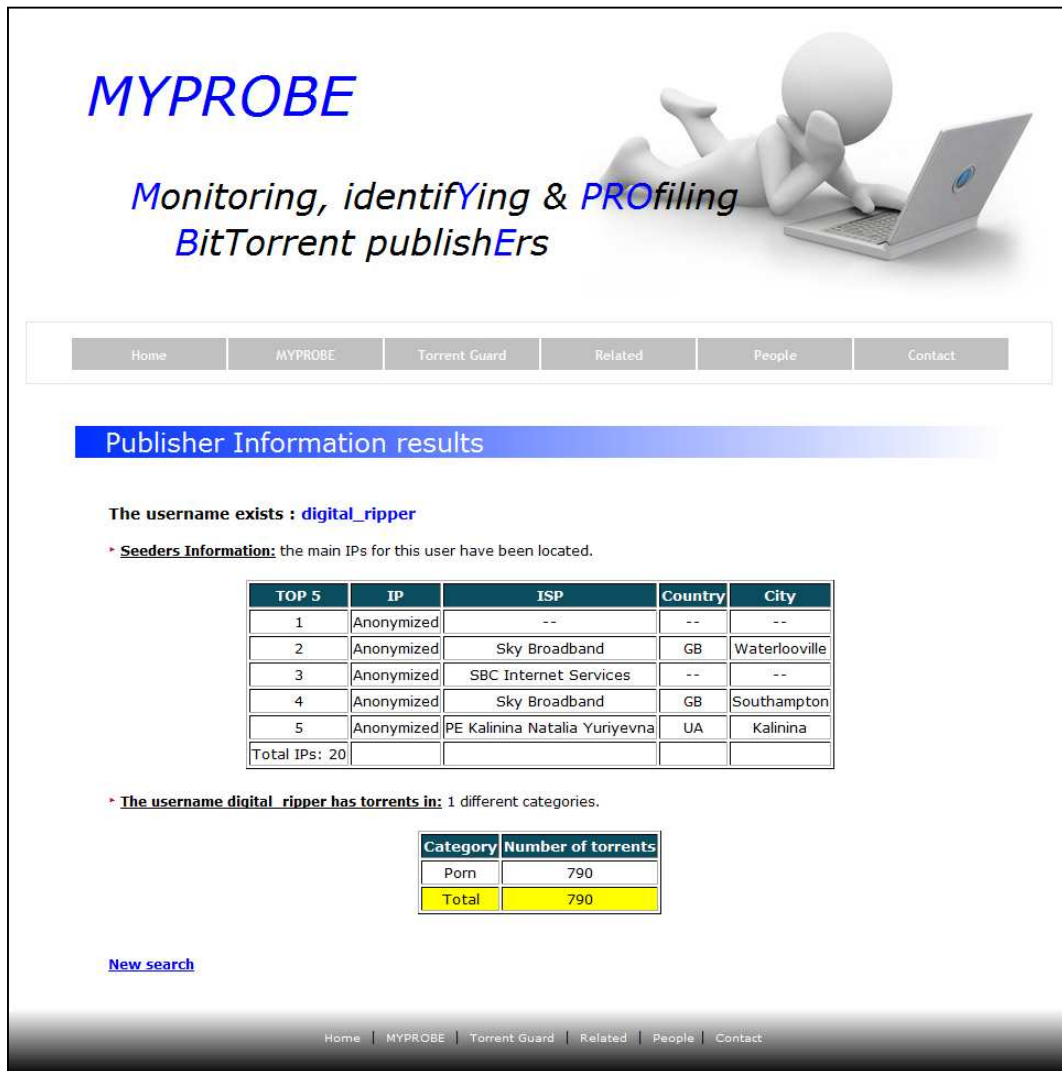
[New search](#)

Home | MYPROBE | Torrent Guard | Related | People | Contact

Figura 22. Resultado consulta de un usuario concreto (TvTeam) en Publishers Information.

Como se deduce del resultado mostrado en la figura anterior, **TvTeam** publica torrents desde más de 1000 direcciones IP distintas. Además, observando el detalle de las 5 primeras direcciones, se advierte que los lugares donde están ubicados los ISP son muy dispares. Por lo tanto, este usuario podría haber contratado distintos servidores de alojamiento web para llevar a cabo su actividad. En este caso, también se aprecia que **TvTeam** comparte más de 3400 archivos *.torrent* dentro de la categoría de Video, no mostrando actividad dentro de otra categoría.

Ahora, se introduce otro nombre de usuario detectado: **digital_ripper**. Dicho usuario suele estar situado dentro de las 10 primeras posiciones de los principales publicadores de contenido en The Pirate Bay.



MYPROBE
Monitoring, identifying & PROfiling
BitTorrent publishErs

Home MYPROBE Torrent Guard Related People Contact

Publisher Information results

The username exists : [digital_ripper](#)

• **Seeders Information:** the main IPs for this user have been located.

TOP 5	IP	ISP	Country	City
1	Anonymized	--	--	--
2	Anonymized	Sky Broadband	GB	Waterlooville
3	Anonymized	SBC Internet Services	--	--
4	Anonymized	Sky Broadband	GB	Southampton
5	Anonymized	PE Kalina Natalia Yuriyevna	UA	Kalina
Total IPs: 20				

• The username [digital_ripper](#) has torrents in: 1 different categories.

Category	Number of torrents
Porn	790
Total	790

[New search](#)

Home | MYPROBE | Torrent Guard | Related | People | Contact

Figura 23. Resultado consulta de un usuario concreto (digital_ripper) en Publishers Information.

Por un lado, como sucedía con **TvTeam**, se muestra la información sobre los principales seeders asociados con el nombre de usuario. En este caso, se resumen las 5 primeras direcciones de las 20 totales. De esta forma se pueden conocer las direcciones IP que han utilizado para publicar contenido. A continuación se muestra el proveedor de servicio de Internet utilizado así como su localización. **Digital_ripper** ha compartido 790 .torrent hasta la fecha rastreados por la aplicación.

Se puede plantear un escenario distinto en el cual mostrar al usuario información sobre el número de torrents y categorías en las que publica un usuario concreto. Sin embargo, no es viable ofrecer información sobre las direcciones IP a través de las cuales realiza esta actividad.

6.2.2 TOP PUBLISHERS

Ahora se inicia la consulta de los principales publicadores accediendo desde el enlace situado en el apartado MYPROBE. Como se observa, el usuario tiene diferentes enlaces a través de los cuales puede iniciar la consulta acerca de los principales publicadores de contenido dentro de BitTorrent.



Figura 24. Acceso a TOP Publishers.

En este punto, si no se indica ningún criterio de consulta y se envía la petición, la página web muestra el siguiente mensaje.



Figura 25. Mensaje de ayuda al usuario en TOP Publishers.

En este caso, se permite al usuario introducir un número entre 1 y 100. Por este motivo solo se pueden introducir tres caracteres dentro del cuadro de diálogo. La página web también es capaz de detectar si se están introduciendo caracteres alfanuméricos. De esta forma si se introducen letras, por ejemplo: **aaa** se muestra el mensaje.

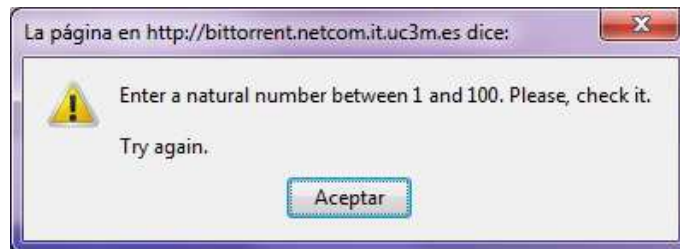


Figura 26. Mensaje de ayuda al usuario en TOP Publishers cuando se introducen valores no permitidos.

Si se combinan caracteres alfanuméricos no permitidos y además se filtra por alguna de las categorías del panel desplegable, se muestra el mismo mensaje de la figura anterior.

Únicamente al introducir un valor correcto, entre 1 y 100, se obtiene un resultado satisfactorio.

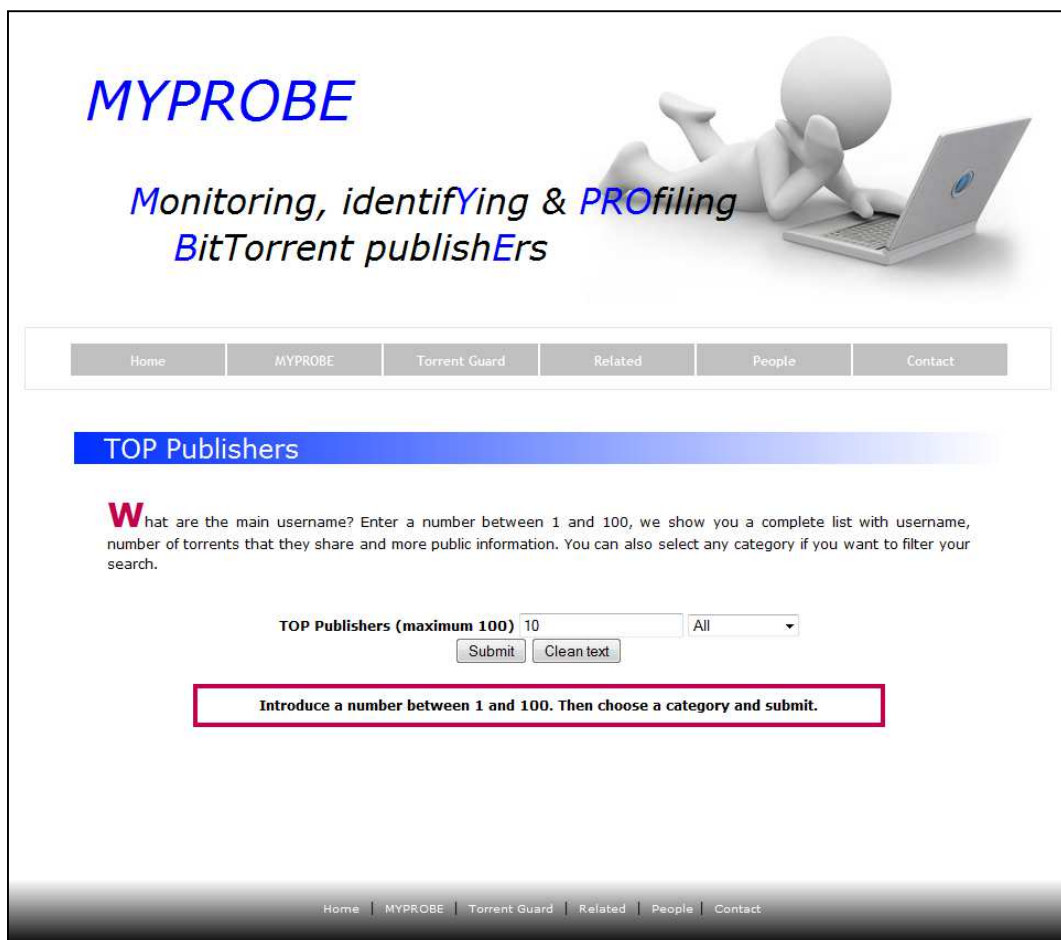
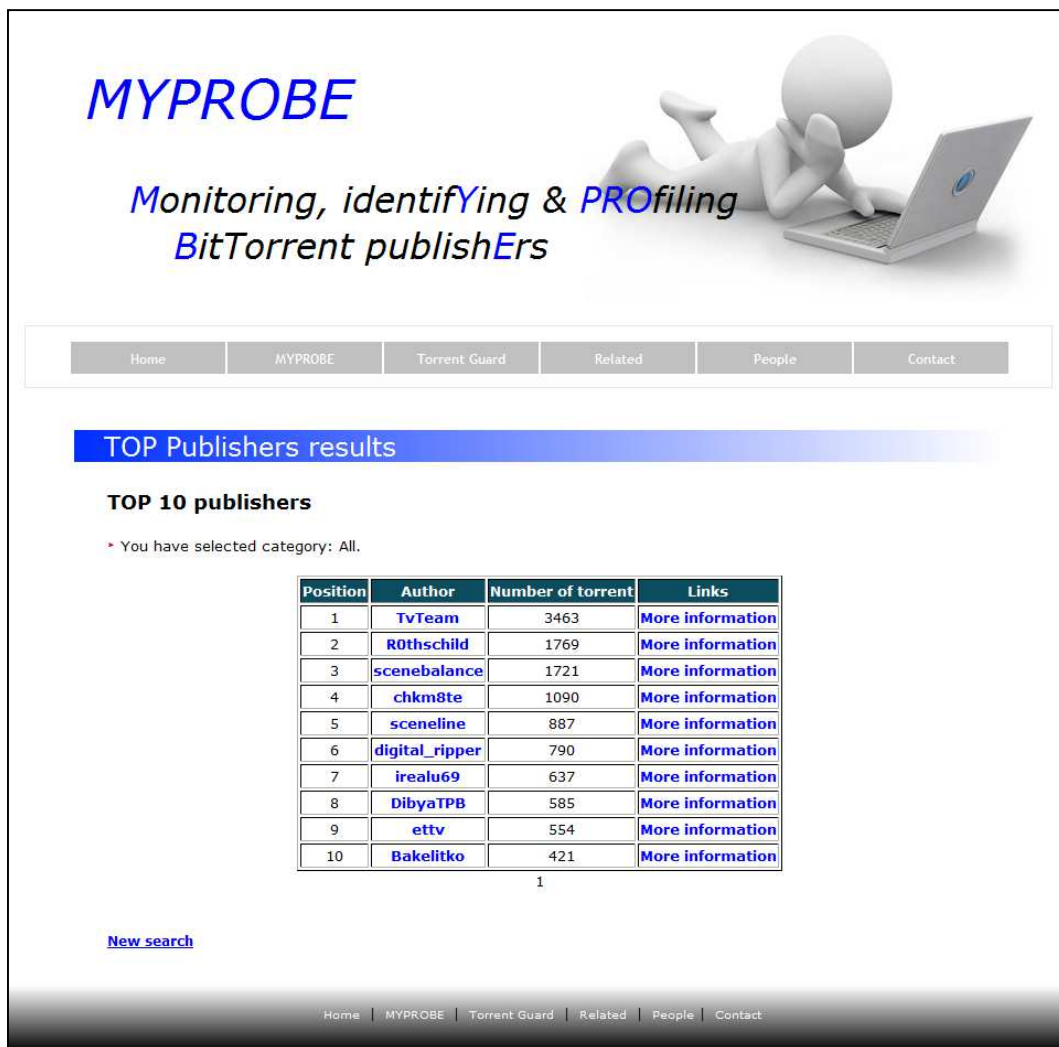


Figura 27. Consulta correcta en TOP Publishers.

En la Figura 27 se muestra la apariencia que tiene el apartado TOP Publishers en el portal web. En este caso se ha optado por consultar los **10** primeros publicadores independientemente de la categoría.

Como se aprecia en la figura anterior, en todo momento se indica al usuario lo que debe hacer para completar la consulta de forma satisfactoria. El mensaje dentro del recuadro magenta aparece cuando se va a rellenar el formulario de consulta. Éste desaparece cuando el ratón se sitúa fuera del formulario. La página web se autoajusta para poder mostrar el mensaje correctamente.



MYPROBE

*Monitoring, identifyIng & PROfiling
BitTorrent publishErs*

Home MYPROBE Torrent Guard Related People Contact

TOP Publishers results

TOP 10 publishers

You have selected category: All.

Position	Author	Number of torrent	Links
1	TvTeam	3463	More information
2	R0thschild	1769	More information
3	scenebalance	1721	More information
4	chkm8te	1090	More information
5	sceneline	887	More information
6	digital_ripper	790	More information
7	irealu69	637	More information
8	DibyaTPB	585	More information
9	etv	554	More information
10	Bakelitko	421	More information

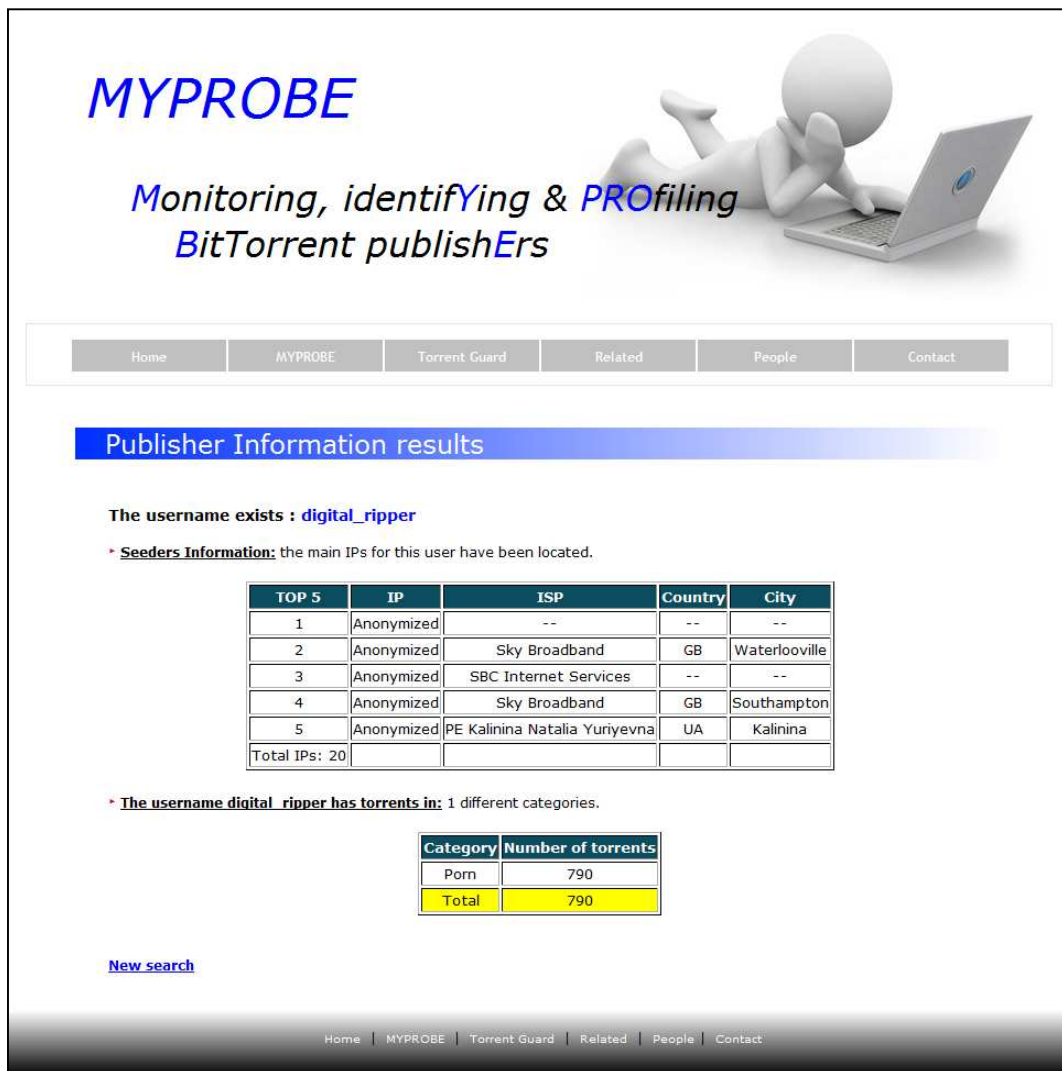
[New search](#)

Home MYPROBE Torrent Guard Related People Contact

Figura 28. Resultado consulta correcta en TOP Publishers.

Al introducir el valor **10** y enviar la petición la página web muestra el resultado. En primer lugar se muestra un pequeño título que resume la información consultada. En este caso: *TOP 10 publishers*. *You have selected category: All*. Es decir, los 10 primeros nombres de usuario que comparten mayor número de torrents a través del portal web The Pirate Bay. Al no indicar ningún filtro en la categoría, muestra la información independientemente del tipo de *.torrent*

que sea. La información se representa en una tabla organizando el resultado en función del número de torrents de mayor a menor. Además, por cada usuario se añade un enlace directo para conocer más información. Si, por ejemplo, se pulsa el enlace de *More information* para **digital_ripper**, se mostraría la misma información que al introducir el nombre de usuario a través del enlace Publishers Information.



MYPROBE
Monitoring, identifyIng & PROfiling
BitTorrent publishErs

Home MYPROBE Torrent Guard Related People Contact

Publisher Information results

The username exists : **digital_ripper**

• **Seeders Information:** the main IPs for this user have been located.

TOP 5	IP	ISP	Country	City
1	Anonymized	--	--	--
2	Anonymized	Sky Broadband	GB	Waterlooville
3	Anonymized	SBC Internet Services	--	--
4	Anonymized	Sky Broadband	GB	Southampton
5	Anonymized	PE Kalina Natalia Yuriyevna	UA	Kalina
Total IPs: 20				

• The username **digital_ripper** has torrents in: 1 different categories.

Category	Number of torrents
Porn	790
Total	790

[New search](#)

Home | MYPROBE | Torrent Guard | Related | People | Contact

Figura 29. Resultado del enlace More Information en TOP Publishers para digital_ripper.

6.2.3 TOP ISP

El objetivo es mostrar los principales proveedores de servicio de Internet que utilizan los usuarios de The Pirate Bay para publicar torrents.

El funcionamiento es idéntico al explicado en el apartado anterior para consultar los principales publicadores: TOP Publishers. Se puede indicar el número de ISP principales que se quiere ver. Así como indicar alguna de las categorías que aparecen en la lista desplegable. Este dato no es obligatorio. Igualmente, a través de mensajes al usuario, se guía al mismo en caso de producirse algún error en la consulta.

Por ejemplo, si se accede a TOP ISP desde el menú principal y, sin seleccionar nada, se manda la consulta, aparecerá la siguiente ventana:



Figura 30. Mensaje de ayuda al usuario en TOP ISP.

El número de caracteres está limitado a tres, ya que no es posible indicar un número mayor a 100. En caso de introducir: letras, caracteres o una combinación de ambos, filtrando o no para una categoría concreta, la página detecta el error y muestra el siguiente mensaje:

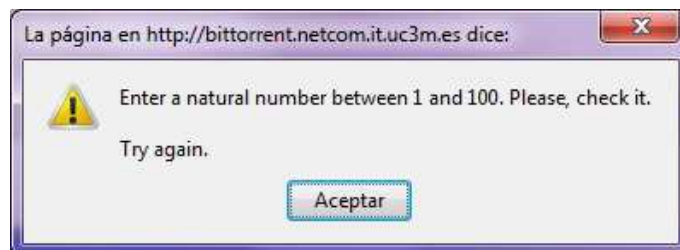
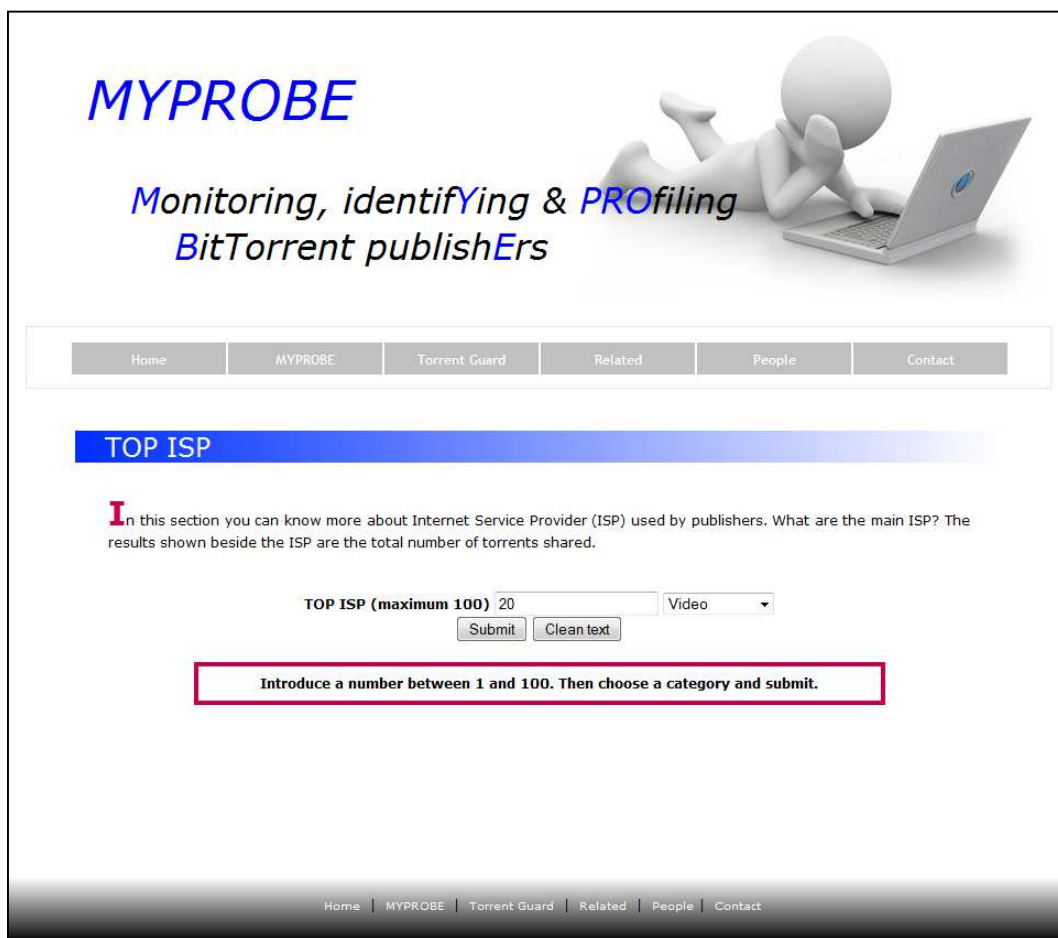


Figura 31. Mensaje de ayuda al usuario en TOP ISP cuando se introducen valores no permitidos.

Finalmente, si se introduce un valor entre 1 y 100 la página reconoce que el valor es correcto y muestra el resultado correspondiente a la consulta. Las figuras siguientes muestran el resultado:



The screenshot displays the MYPROBE website interface. At the top, the logo "MYPROBE" is in blue, followed by the tagline "Monitoring, identifying & PROFiling BitTorrent publishers" in a mix of blue and black italicized fonts. To the right is a 3D illustration of a person lying on their back, propped up on one arm, using a laptop. Below the header is a navigation bar with links: Home, MYPROBE, Torrent Guard, Related, People, and Contact. The main section is titled "TOP ISP" in a blue gradient box. Below this, a paragraph explains the section's purpose: "In this section you can know more about Internet Service Provider (ISP) used by publishers. What are the main ISP? The results shown beside the ISP are the total number of torrents shared." The search form includes a label "TOP ISP (maximum 100)", a text input field containing "20", and a dropdown menu set to "Video". There are "Submit" and "Clean text" buttons. A red-bordered box contains the instruction: "Introduce a number between 1 and 100. Then choose a category and submit." The footer repeats the navigation links.

MYPROBE

*Monitoring, identifying & PROFiling
BitTorrent publishers*

Home MYPROBE Torrent Guard Related People Contact

TOP ISP

In this section you can know more about Internet Service Provider (ISP) used by publishers. What are the main ISP? The results shown beside the ISP are the total number of torrents shared.

TOP ISP (maximum 100) 20 Video

Submit Clean text


Introduce a number between 1 and 100. Then choose a category and submit.

Home MYPROBE Torrent Guard Related People Contact

Figura 32. Consulta correcta en TOP ISP.

MYPROBE

*Monitoring, identifying & PROfiling
BitTorrent publishErs*



[Home](#) | [MYPROBE](#) | [Torrent Guard](#) | [Related](#) | [People](#) | [Contact](#)

TOP ISP results

TOP 20 ISP

• You have selected category: Video.

• Page 1 of 2 :

Position	ISP	Country	Number of torrents
1	Comcast Cable	US	358
2	Road Runner	US	280
3	Ovh Systems	FR	224
4	Shaw Communications	CA	100
5	Verizon Internet Services	US	95
6	Telefonica de Espana	ES	92
7	Cox Communications	US	84
8	Com Hem AB	SE	79
9	TeliaSonera AB	SE	74
10	Casema B.V.	NL	73


1 ...>> [Next page 2](#)

• You can see [all the results](#).

[New search](#)

[Home](#) | [MYPROBE](#) | [Torrent Guard](#) | [Related](#) | [People](#) | [Contact](#)

Figura 33. Resultado consulta correcta en TOP ISP primera tabla.



MYPROBE

*Monitoring, identifying & profiling
BitTorrent publishers*

Home | MYPROBE | Torrent Guard | Related | People | Contact

TOP ISP results

TOP 20 ISP

* You have selected category: Video.

* Page 2 of 2 :

Position	ISP	Country	Number of torrents
11	Telstra Internet	AU	69
12	Charter Communications	US	69
13	LeaseWeb B.V.	NL	68
14	Bredbandsbolaget AB	SE	63
15	SBC Internet Services	US	62
16	Sociedad Promotora de las	ES	60
17	Cableuropa S.A.U.	ES	57
18	Embratel	BR	56
19	Global Village Telecom LTDA.	BR	47
20	Chello	NL	46

[Previous page 1](#) <<... 2

* You can see [all the results](#).

[New search](#)

Home | MYPROBE | Torrent Guard | Related | People | Contact

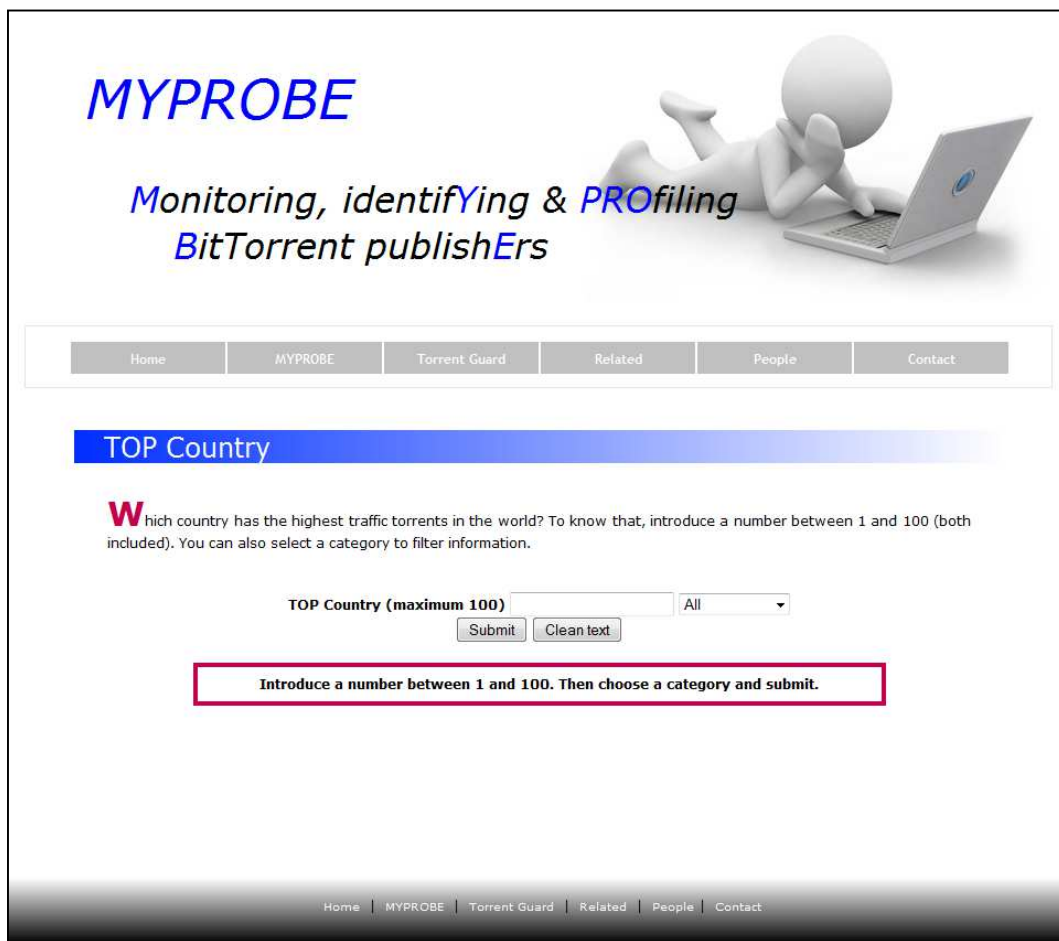
Figura 34. Resultado consulta correcta en TOP ISP segunda tabla.

El resultado se muestra enmarcado en una tabla que tiene como título el resumen de la consulta. En este caso: *TOP 20 ISP. You have selected category: Video.* Se puede observar en la Figura 32 que los valores introducidos han sido **20** y **Video**.

6.2.4 TOP COUNTRY

Gracias a la herramienta GeoIP proporcionada por MaxMind, Crawler es capaz de identificar el ISP, el país y ciudad del *.torrent* publicado. Por lo tanto, la base de datos que consulta la página web contiene esta información que es interesante para los usuarios. De esta manera se puede saber donde están ubicados los ISP, aunque esto no significa que los propios usuarios estén localizados en el mismo lugar.

Se accede al enlace TOP Country del menú principal. En la Figura 35 se muestra el aspecto inicial de la página de consulta:

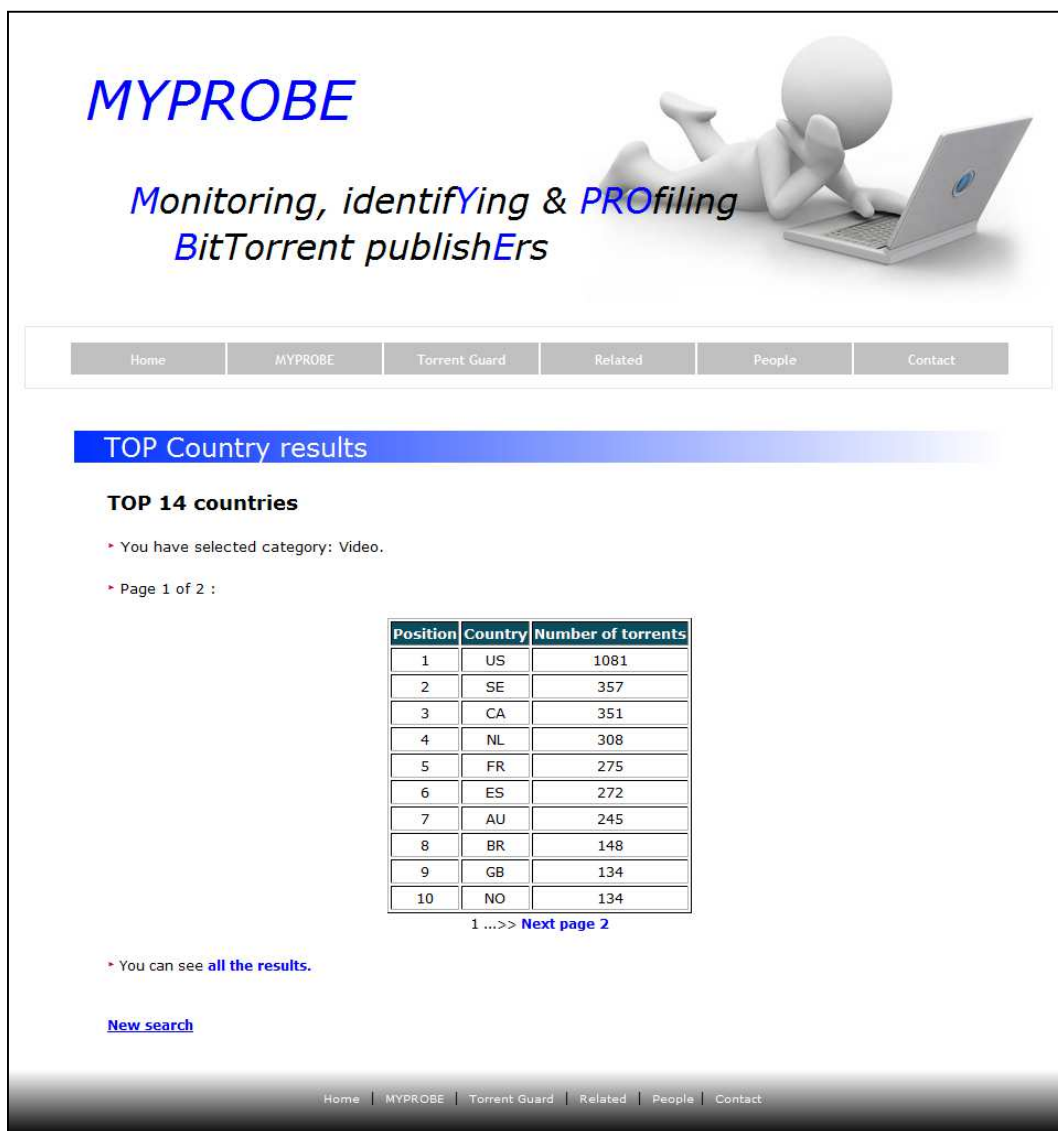


The screenshot displays the MYPROBE website interface. At the top, the logo "MYPROBE" is in blue, followed by the tagline "Monitoring, identifying & Profiling BitTorrent publishers" in a mix of blue and black text. To the right is a 3D illustration of a person lying down using a laptop. Below the header is a navigation menu with links: Home, MYPROBE, Torrent Guard, Related, People, and Contact. The main section is titled "TOP Country" in a blue gradient box. Below this, a paragraph explains the search: "Which country has the highest traffic torrents in the world? To know that, introduce a number between 1 and 100 (both included). You can also select a category to filter information." The search form includes a text input labeled "TOP Country (maximum 100)", a dropdown menu set to "All", and "Submit" and "Clean text" buttons. A red-bordered box contains the instruction: "Introduce a number between 1 and 100. Then choose a category and submit." The footer repeats the navigation links.

Figura 35. Acceso a TOP Country.

El comportamiento de la página es el mismo que en los casos anteriores para: TOP Publishers y TOP ISP. Si el usuario no introduce un número entre 1 y 100 y envía la consulta, aparecerá una ventana emergente al comprobar que no se ha introducido ningún valor (como sucedía en los casos anteriores, ver Figura 25 o Figura 30). Si se introduce algún tipo de carácter no válido, la página web mostrará una nueva ventana emergente con el mismo mensaje que en casos anteriores (ver Figura 26 y Figura 31).

Se prueba la página indicando como valores de la consulta: **14** y categoría **Video**. El resultado se muestra en la figura siguiente:



MYPROBE

*Monitoring, identifyIng & PROfiling
BitTorrent publishErs*

Home MYPROBE Torrent Guard Related People Contact

TOP Country results

TOP 14 countries

► You have selected category: Video.

► Page 1 of 2 :

Position	Country	Number of torrents
1	US	1081
2	SE	357
3	CA	351
4	NL	308
5	FR	275
6	ES	272
7	AU	245
8	BR	148
9	GB	134
10	NO	134

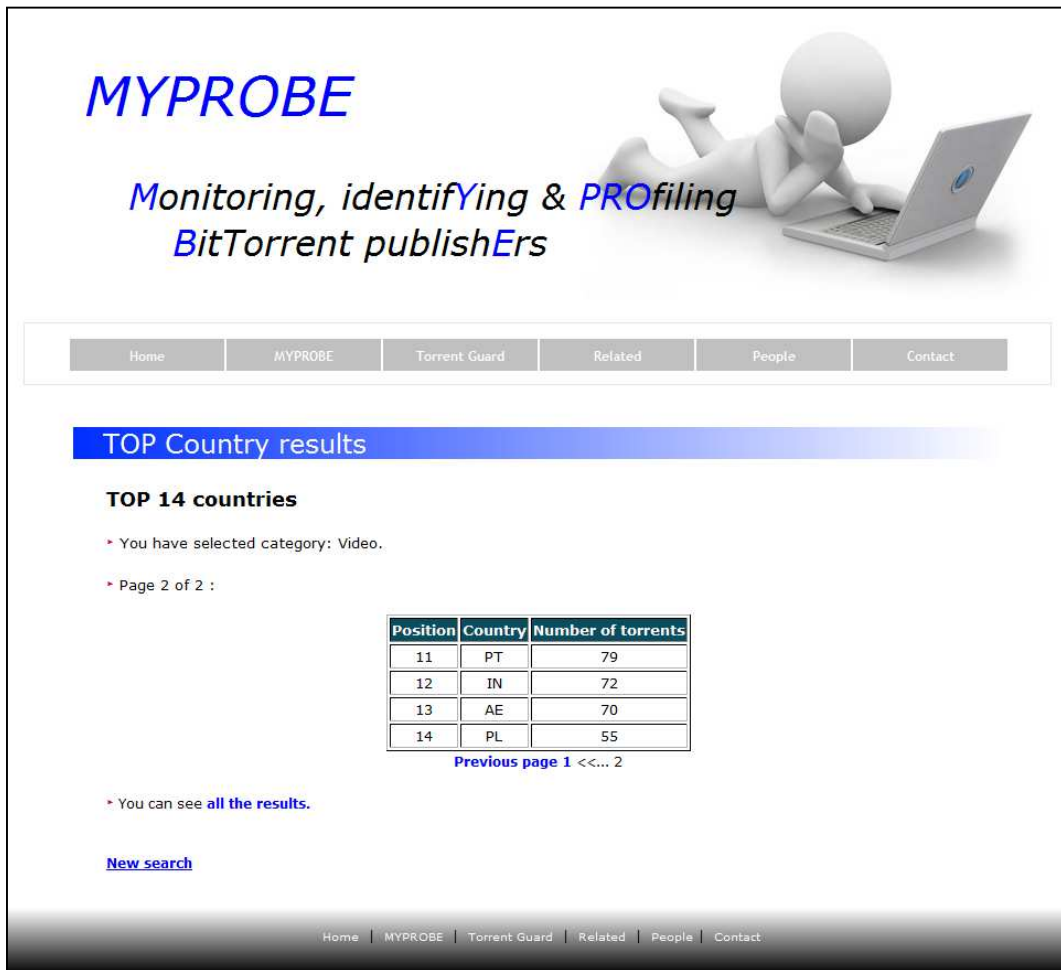
1 ...>> [Next page 2](#)

► You can see [all the results](#).

[New search](#)

Home MYPROBE Torrent Guard Related People Contact

Figura 36. Resultado consulta correcta en TOP Country primera tabla.



The screenshot shows the MYPROBE website interface. At the top, the logo 'MYPROBE' is displayed in blue, followed by the tagline 'Monitoring, identifyIng & PROfiling BitTorrent publishErs' in a mix of blue and black text. To the right, a 3D white figure is lying on its back, using a laptop. Below the header is a navigation bar with links: Home, MYPROBE, Torrent Guard, Related, People, and Contact. The main content area features a blue gradient bar with the text 'TOP Country results'. Below this, the section 'TOP 14 countries' is shown. A message indicates 'You have selected category: Video.' and 'Page 2 of 2 :'. A table displays the top 14 countries by the number of torrents. The table has three columns: Position, Country, and Number of torrents. The data rows show positions 11 to 14 for countries PT, IN, AE, and PL with 79, 72, 70, and 55 torrents respectively. Below the table, there is a link 'Previous page 1 <<... 2'. A message states 'You can see all the results.' and a link 'New search' is provided. At the bottom, a footer navigation bar repeats the links: Home, MYPROBE, Torrent Guard, Related, People, and Contact.

Position	Country	Number of torrents
11	PT	79
12	IN	72
13	AE	70
14	PL	55

Figura 37. Resultado consulta correcta en TOP Country segunda tabla.

En este caso se ha obtenido en primer lugar Estados Unidos con un total de 1081 torrents publicados en la categoría de **Video**. Como se ha comentado antes, puede que los usuarios que han publicado esos torrents no estén ubicados físicamente en estos países. Pero sí significa que un usuario está utilizando los recursos ofrecidos por un proveedor de servicio de Internet alojados en este país.

6.3 CONSULTAS REALIZADAS EN TORRENT GUARD

En segundo lugar se van a analizar los resultados obtenidos tras consultar información sobre distintos archivos *.torrent*. De esta forma, se contemplan los distintos escenarios y muestran los distintos mensajes que aparecen a través de las páginas web desarrolladas.

6.3.1 CHECK YOUR TORRENT

Una de las opciones para acceder al apartado de consulta de torrents está situado en el menú principal: Check your torrent.

Torrent Guard

Home MYPROBE Torrent Guard Related People Contact

Check your torrent

Do you want to know whether your torrent is malware or not? In this section you can check it. Even if you have a torrent file downloaded or not, we give you three different options to check the status of it. Insert the information you have and we take care to identify if your torrent is malware or not. Free, easy and fast!

- Before starting to download a file, be sure about its status. Select a torrent file of your computer, in a few seconds we tell you what is the result.

Select Torrent:

- If you rather check a file before downloading it, don't worry we provide you the chance to do it! It's quite easy you just need to copy hash information of the torrent and paste in this box.

Introduce infohash:

- Taking into account the latest develop in torrent environment, you can also check a magnet link. It's as easy as before: copy, paste and submit!

Introduce magnet link:

Home MYPROBE Torrent Guard Related People Contact

Figura 38. Acceso a Check your torrent.

Una vez situados en este subapartado, la página ofrece al usuario tres opciones distintas para comprobar un *.torrent*: añadir directamente un archivo *.torrent* ubicado en algún

directorio personal, teclear el valor de información hash asociada a un archivo o introducir un enlace magnético. Al estar todas las opciones ubicadas en el mismo subapartado se da a los usuarios la opción de elegir la forma de consulta que mejor se adapte a sus necesidades.

A continuación se van a analizar los distintos resultados que muestra la página en función de la opción elegida, de esta forma se pretende recorrer todos los escenarios de consulta/respuesta posibles.

6.3.1.1 OPCIÓN 1: COMPROBAR ARCHIVO .TORRENT

Para esta primera opción, se va a analizar cómo responde la página web en función de las distintas consultas. Por un lado puede que el usuario, de forma intencionada o no, intente enviar una consulta sin introducir ningún archivo o adjuntando uno que no sea un *.torrent*. Por otro lado, si el archivo añadido es el adecuado la página muestra un resultado distinto en función del estado del archivo.

En primer lugar, si no se añade ningún archivo y se envía la solicitud de consulta, en la página web aparece el siguiente mensaje de la Figura 39.

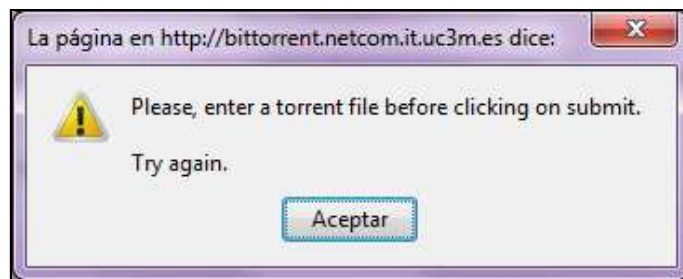


Figura 39. Mensaje de ayuda al usuario en la opción 1 de Check your torrent.

Ahora se introduce un archivo cualquiera dentro de la primera opción que aparece en Check your torrent: **enlaces.txt**. En este caso se trata de un archivo personal que no ha sido descargado de ninguna página web concreta. El archivo en cuestión es un fichero de texto que contiene algunos enlaces en su interior. Al enviar la consulta se obtiene el siguiente resultado:

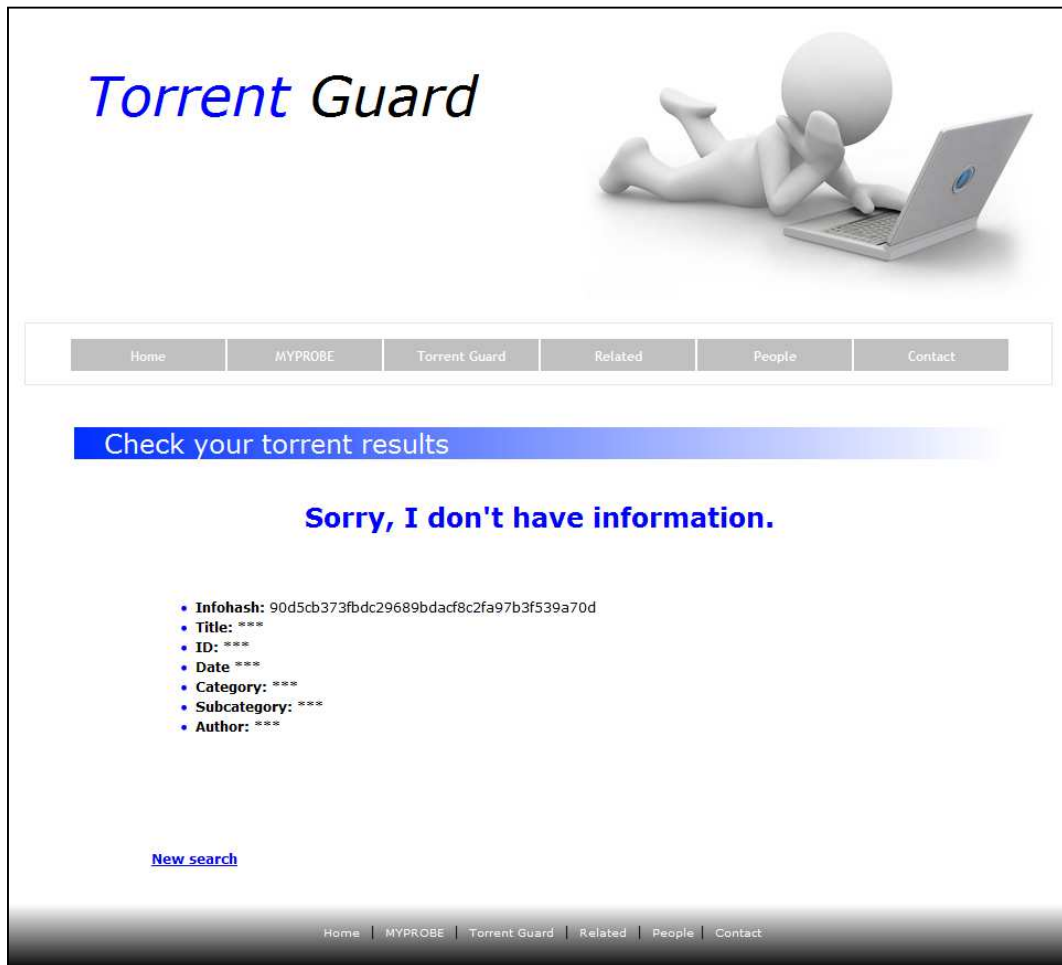


Figura 40. Resultado consulta archivo enlaces.txt.

En primer lugar se indica al usuario que no se dispone de información para valorar si el archivo es falso o no. Esta respuesta es estándar, se ha decidido mostrar este resultado para todos aquellos archivos sobre los cuales no se puede valorar su estado.

Este escenario se repite en caso de consultar un archivo *.torrent* del cual no se tiene demasiada información todavía. Para entender mejor lo que sucede, se repite la consulta para el archivo **M_Clan_-_Retrovision_(1995-2006)_[2006]_-_Pop.3496538.TPB.torrent**.



Figura 41. Resultado consulta archivo M_Clan_-_Retrovision_(1995-2006)_[2006]_-_Pop.3496538.TPB.torrent.

Por lo tanto, en este escenario al usuario simplemente se le indica que no se tiene información sobre el archivo en cuestión. Se muestra por pantalla la información hash del archivo y el resto de valores aparecen en blanco.

Este resultado implica que Torrent Guard no ha identificado el resultado como malware o como correcto (no malware). Por lo tanto, se da una respuesta algo ambigua al usuario ya que el sistema de detección no ha podido clasificar el archivo .torrent dentro de un perfil más concreto. Es posible que si se realiza la consulta sobre el mismo archivo pasado cierto tiempo, la respuesta pueda variar.

Se analiza ahora el archivo **Law.and.Order.SVU.S14E06.HDTV.x264-LOL.7779386.TPB**, el cual ha sido descargado desde The Pirate Bay.



[Home](#) | [MYPROBE](#) | [Torrent Guard](#) | [Related](#) | [People](#) | [Contact](#)

Check your torrent results

PERFECT! TRUSTED CONTENT

- **Title:** Law.and.Order.SVU.S14E06.HDTV.x264-LOL
- **ID:** 7779386
- **Date:** 2012-11-01 02:08:05
- **Category:** Video
- **Subcategory:** TV shows
- **Author:**
[TvTeam](#)
- **Comment:** Torrent downloaded from <http://thepiratebay.se>
- **Piece length:** 262144
- **Hash info:** bc1d09aeb189e069177cba1365961c3efe05d57
- **Hash info encode:** bc1d09aeb189e069177cba1365961c3efe05d57
- **Hash info byte:** %BC%1D%09%AE%CB%18%9E%06%91w%CB%A16Ya%C3%EF%E0%5DW

▸ **Seeders Information:** the main IPs for this user have been located.

TOP5	IP	ISP	Country	City
1	Anonymized	SK Broadband Co Ltd	We don't know	We don't know
2	Anonymized	Road Runner	US	Westminster
3	Anonymized	Bredbandsbolaget AB	SE	Landskrona
4	Anonymized	University of Tromso, Norway	NO	Tromsø
5	Anonymized	Com Hem AB	SE	Stockholm
Total IPs: 1185				

▸ **More Information**

- [User homepage](#)
- [Torrent homepage](#)

[New search](#)

[Home](#) | [MYPROBE](#) | [Torrent Guard](#) | [Related](#) | [People](#) | [Contact](#)

Figura 42. Resultado consulta archivo Law.and.Order.SVU.S14E06.HDTV.x264-LOL.7779386.TPB.

En primer lugar se indica al usuario que el archivo *.torrent* no es malware. Esto es posible ya que la herramienta Torrent Guard ha analizado el archivo buscando alguna coincidencia dentro de sus bases de datos. A continuación, se muestra información básica relacionada con el archivo: el nombre, el identificador que tiene, la fecha de creación, categoría y subcategoría a la que pertenece, el autor del archivo *.torrent* dentro de The Pirate Bay, algún comentario que contenga el archivo, el tamaño de las piezas en las que se divide y la información hash en distintos formatos.

Esta es una información de gran utilidad para los usuarios que desean descargar el archivo. En primer lugar porque se indica de forma clara el estado del archivo y en segundo lugar, porque se muestra de forma esquemática la información relevante sobre el archivo.

A continuación, se resume en una pequeña tabla más información relacionada con el autor del archivo. Como sucedía en el apartado Publisher Information, esta tabla indica a través de qué ISP el autor ha publicado otros archivos. En este ejemplo se observa que *TvTeam* ha publicado desde 1185 direcciones IP distintas, mostrando en la tabla las 5 primeras. En caso de que el autor hubiera publicado desde menos de 5 direcciones, la tabla resumen se adapta al número total de IPs utilizadas.

Por último, en caso de que el usuario quiera obtener más información acerca del archivo *.torrent* o sobre el autor, se han añadido dos enlaces: el primero, redirige a la página personal que el autor tiene en The Pirate Bay y, el segundo, da acceso a la página que tiene el archivo *.torrent* dentro del mismo portal de descargas.

Se analiza ahora el estado del archivo *.torrent* **Cubase_SX3_Magyar.oktat_oacute_-DVD.7639015.TPB**, el cual también ha sido descargado de The Pirate Bay. Como se observa en la Figura 43, se trata de un archivo falso. Esto significa que la herramienta de detección Torrent Guard ha clasificado el archivo como dañino para el usuario. En este caso se reduce la información mostrada acerca del archivo, mostrando simplemente la información más básica del mismo.

Como se explicó en el Capítulo 4, una de las opciones por la que Torrent Guard ha identificado este archivo como falso puede ser que el usuario ha sido borrado del portal The Pirate Bay debido a la publicación fraudulenta a través de una dirección IP (repitiéndose esta estrategia con 3 o más usuarios distintos, para una misma dirección IP). Otro de los motivos que ha llevado a Torrent Guard a clasificar este archivo como malicioso puede ser que conoce, incluso antes que el propio portal de descargas, que el archivo ha sido publicado desde una dirección IP considerada peligrosa.

Por este motivo, no se muestra ninguna información más concreta acerca del username que ha publicado el archivo. Podría darse el caso, para algunos username, de contar con información acerca de las direcciones IP o ISP a través de los cuales publican los archivos pero, como no es posible mostrarla en todos los casos, se ha optado por limitar esta información.

Torrent Guard



[Home](#) | [MYPROBE](#) | [Torrent Guard](#) | [Related](#) | [People](#) | [Contact](#)

Check your torrent results

BE CAREFUL! FAKE CONTENT

- **Title:** Cubase SX3 Magyar.óktató-DVD
- **ID:** 7639015
- **Date:** 2012-09-15 16:07:26
- **Category:** Video
- **Subcategory:** Other
- **Author:**
[stef666stef](#)
- **Comment:** Torrent downloaded from <http://thepiratebay.se>
- **Piece length:** 2097152
- **Hash info:** 5453da50840c7ff7152c36d2c167c74f0cae706a
- **Hash info encode:** 5453da50840c7ff7152c36d2c167c74f0cae706a
- **Hash info byte:** TS%DAP%84%0C%7F%F7%15%2C6%D2%C1g%C7O%0C%AEpj

[New search](#)

[Home](#) | [MYPROBE](#) | [Torrent Guard](#) | [Related](#) | [People](#) | [Contact](#)

Figura 43. Resultado consulta archivo Cubase_SX3_Magyar.óktat_oacute_-DVD.7639015.TPB.

6.3.1.2 OPCIÓN 2: COMPROBAR INFOHASH

A continuación se van a analizar las mismas consultas, para el caso de utilizar la segunda de las opciones que aparecen dentro del subapartado Check your torrent.

En esta segunda opción, si no se introduce la información correctamente, aparecerán distintos mensajes en función del error que se haya producido. Si el usuario pulsa el botón Submit sin introducir ningún valor, se muestra el primer mensaje que aparece en la siguiente figura. Si el valor introducido es demasiado corto, se muestra el segundo mensaje. Finalmente el último mensaje aparece si el valor consultado está mal escrito, es decir, contiene algún dígito fuera del rango de los números permitidos en la información hash.

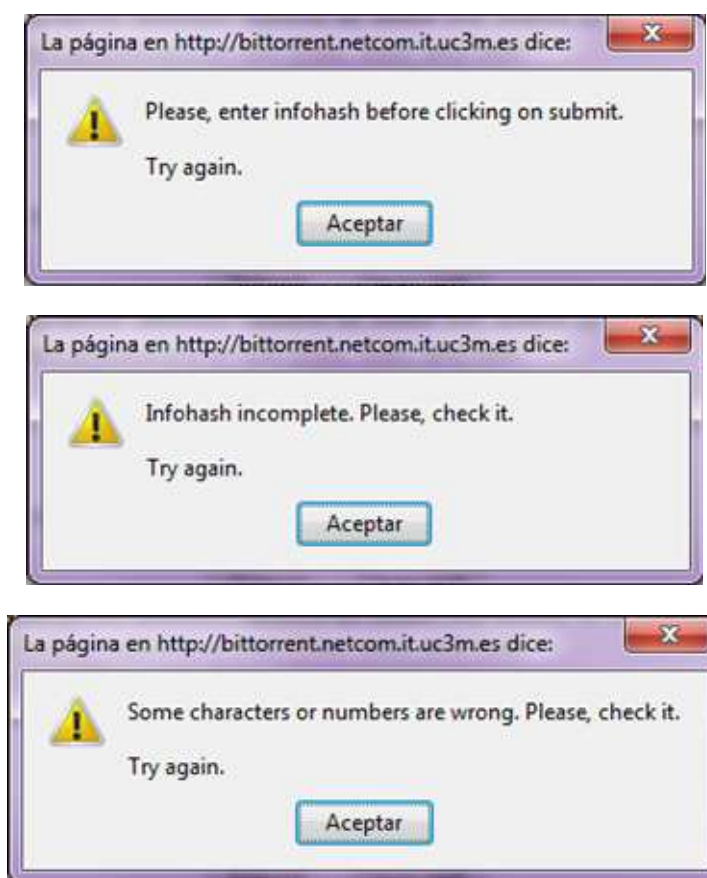


Figura 44. Mensajes de ayuda al usuario en la opción 2 de Check your torrent.

En primer lugar, se ha optado por coger al azar la información hash de uno de los archivos .torrent publicados dentro de The Pirate Bay, este valor es: **65C59094DB4B206AE9B4FC95BB143005DA02BA82.**




Figura 45. Resultado consulta infohash 65C59094DB4B206AE9B4FC95BB143005DA02BA82.

En este caso, como sucedía anteriormente, no se puede clasificar el archivo dentro de un perfil concreto. Por lo tanto se opta por decir al usuario que no se tiene suficiente información. Es posible que una consulta posterior para este archivo tenga como resultado otra respuesta distinta, pero en el momento de la consulta, Torrent Guard no disponía de suficiente información.

Para comprobar el resto de resultados posibles se ha decidido seleccionar el valor de información hash de los anteriores archivos *.torrent* que fueron comprobados para la primera opción.

El primer archivo comprobado, **Law.and.Order.SVU.S14E06.HDTV.x264-LOL.7779386.TPB**, tiene un valor de infohash igual a **bc1d09aecb189e069177cba1365961c3efe05d57**, como se puede ver en la Figura 42. Por lo tanto, éste es el valor introducido en la segunda opción dentro de Check your torrent. El resultado mostrado por la página debe coincidir con el resultado mostrado anteriormente cuando se introducía el archivo dentro de la primera opción. En la Figura 46 se observa que, efectivamente, este resultado coincide con el anterior. Así, queda patente que la página es coherente en los resultados ofrecidos independientemente del formato de la consulta.

Torrent Guard



[Home](#)
[MYPROBE](#)
[Torrent Guard](#)
[Related](#)
[People](#)
[Contact](#)

Check your torrent results

PERFECT! TRUSTED CONTENT

- **Title:** Law.and.Order.SVU.S14E06.HDTV.x264-LOL
- **ID:** 7779386
- **Date:** 2012-11-01 02:08:05
- **Category:** Video
- **Subcategory:** TV shows
- **Author:**
[TvTeam](#)
- **Hash info:** bc1d09aeb189e069177cba1365961c3efe05d57
- **Hash info encode:** bc1d09aeb189e069177cba1365961c3efe05d57
- **Hash info byte:** %BC%1D%09%AE%CB%18%9E%06%91w%CB%A16Ya%C3%EF%E0%5DW

• **Seeders Information:** the main IPs for this user have been located.

TOP5	IP	ISP	Country	City
1	Anonymized	SK Broadband Co Ltd	We don't know	We don't know
2	Anonymized	Road Runner	US	Westminster
3	Anonymized	Bredbandsbolaget AB	SE	Landskrona
4	Anonymized	University of Tromso, Norway	NO	Troms
5	Anonymized	Com Hem AB	SE	Stockholm
Total IPs: 1185				

• **More Information**

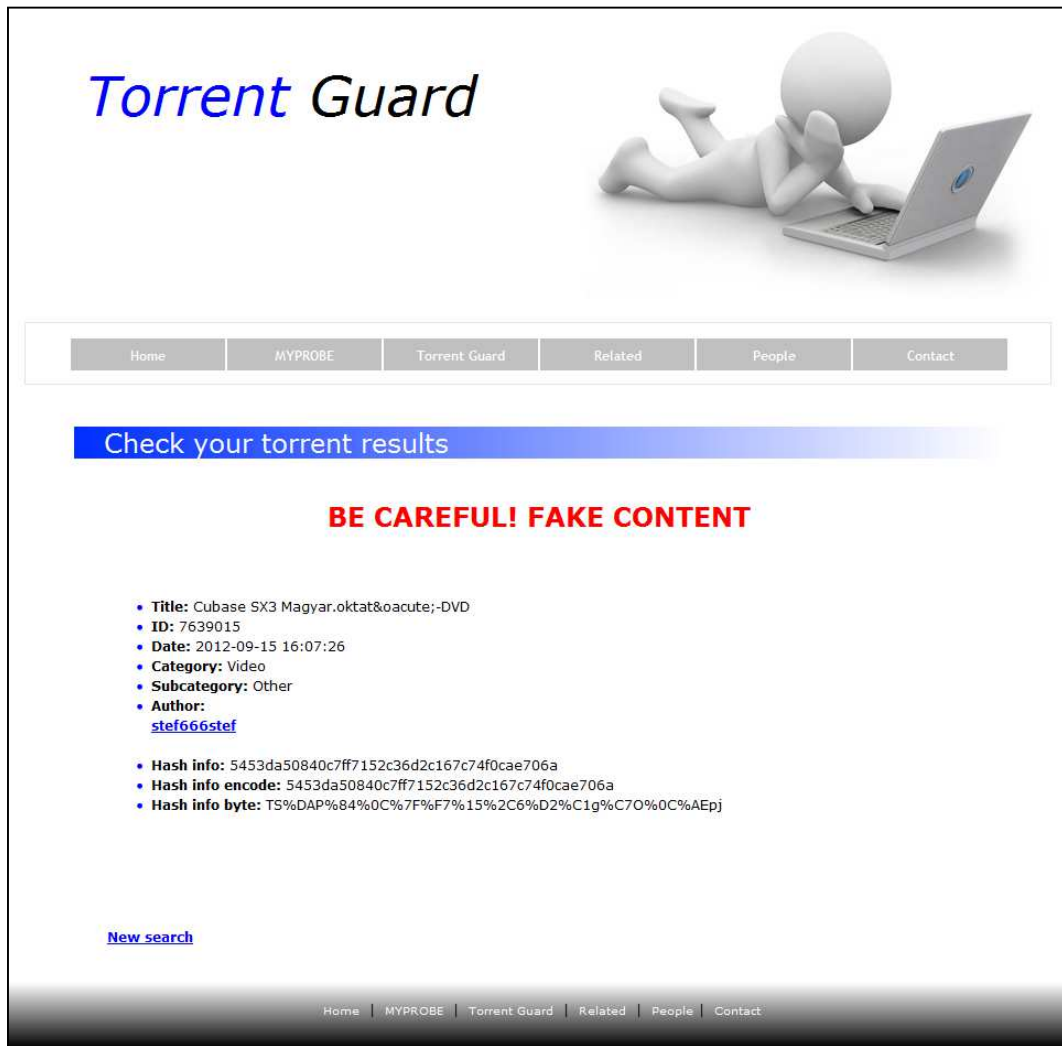
- [User homepage](#)
- [Torrent homepage](#)

[New search](#)

[Home](#)
[MYPROBE](#)
[Torrent Guard](#)
[Related](#)
[People](#)
[Contact](#)

Figura 46. Resultado consulta infohash bc1d09aeb189e069177cba1365961c3efe05d57.

Se procede de la misma manera para comprobar el resultado que ofrece la página si se introduce directamente la información hash del archivo **Cubase_SX3_Magyar.oktat_oacute_-DVD.7639015.TPB**, en este caso sería **5453da50840c7ff7152c36d2c167c74f0cae706a** como se observa en la Figura 47.



Torrent Guard

Home MYPROBE Torrent Guard Related People Contact

Check your torrent results

BE CAREFUL! FAKE CONTENT

- **Title:** Cubase SX3 Magyar.oktató-DVD
- **ID:** 7639015
- **Date:** 2012-09-15 16:07:26
- **Category:** Video
- **Subcategory:** Other
- **Author:** [stef66stef](#)
- **Hash info:** 5453da50840c7ff7152c36d2c167c74f0cae706a
- **Hash info encode:** 5453da50840c7ff7152c36d2c167c74f0cae706a
- **Hash info byte:** TS%DAP%84%0C%7F%F7%15%2C6%D2%C1g%C7O%0C%AEpj

[New search](#)

Home | MYPROBE | Torrent Guard | Related | People | Contact

Figura 47. Resultado consulta infohash 5453da50840c7ff7152c36d2c167c74f0cae706a.

La información hash introducida en la segunda opción dentro de Check your torrent ha generado la misma respuesta dentro de Torrent Guard. Por lo tanto, como sucedía anteriormente, queda demostrada otra de las cualidades de esta aplicación: es capaz de clasificar un archivo simplemente conociendo la información hash. Esto es posible gracias: primero, a que los archivos quedan biunívocamente etiquetados con esta información y, segundo, al análisis que realiza la herramienta y la elaboración de sus bases de datos.

6.3.1.3 OPCIÓN 3: COMPROBAR ENLACE MAGNÉTICO

Para la tercera de las opciones, como sucede cuando se comprueba la infohash, también aparecen distintos mensajes en función del tipo de error que produzcan la cadena de texto introducida. Si se pulsa el botón Submit sin introducir ningún valor, aparece el mensaje: *Por favor, introduce un enlace magnético*. Si se introduce un enlace que no contenga infohash o que no tenga el tamaño adecuado, se muestra el mensaje por pantalla: *Enlace magnético incompleto. Por favor, revise el campo*.

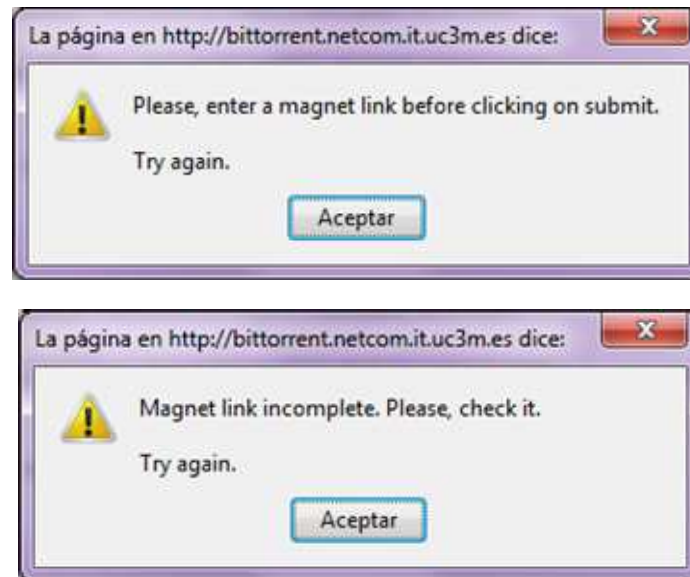



Figura 48. Mensajes de ayuda al usuario en opción 3 de Check your torrent.

Para conseguir un enlace magnético basta con elegir el archivo que se quiere comprobar del portal de descargas The Pirate Bay y situar el ratón sobre el icono . Ahora, pulsando el botón derecho del ratón, se elige la opción Copiar la ruta del enlace. Finalmente se pega el enlace dentro de la tercera opción que aparece en Check your torrent.

En caso de que el enlace copiado no contenga la información hash necesaria, la propia página informará al usuario de que el enlace añadido está incompleto, mostrando el segundo mensaje que aparece en la figura anterior.

En primer lugar se elige al azar un archivo dentro de The Pirate Bay, se trata de: **Shrek Thriller Night 2011 BluRay 1080p DTS x264-3Li**. Al comprobar el estado de este enlace se obtiene el siguiente resultado:

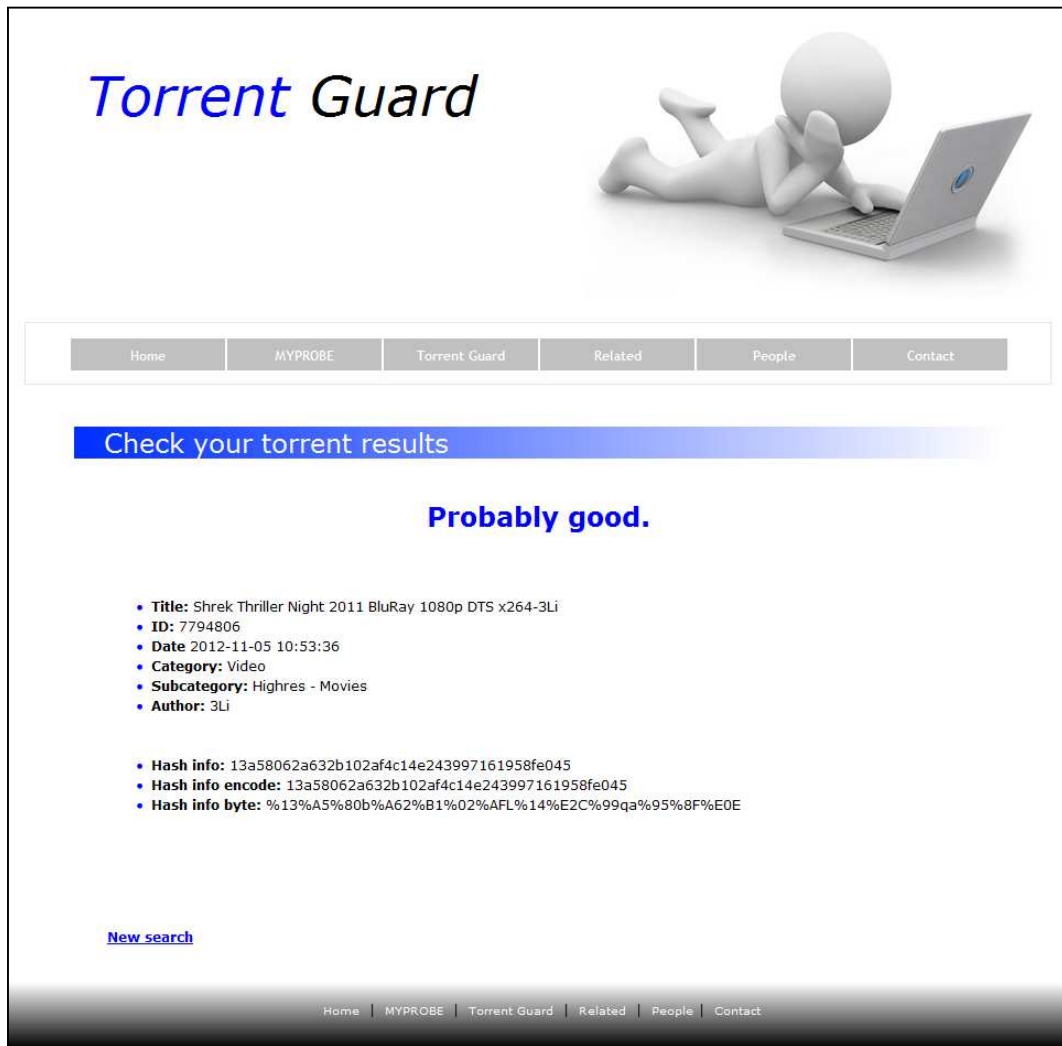


Figura 49. Resultado consulta enlace magnético probablemente correcto.

Este resultado implica que la aplicación sí tiene algún tipo de registro o información dentro de las bases de datos que consulta, pero no puede adelantar si se trata de un archivo malintencionado o no. Se ha optado por mostrar la información más básica del archivo *.torrent*.

Cabe destacar que este archivo, como puede suceder con otros, sí está registrado dentro de las bases de datos de Torrent Guard. Por tanto, podría suceder que al comprobar este mismo enlace en otro momento posterior el resultado varíe clasificando el enlace como malicioso.

A continuación se muestra el aspecto de la página cuando se introduce un enlace magnético cuyo estado no se puede clasificar. Es decir, cuando el resultado de la consulta no permite ofrecer al usuario información relevante sobre el archivo, ni su clasificación dentro de un perfil concreto dentro de Torrent Guard.

El enlace elegido, el cual incluye la información hash del archivo elegido, es: **magnet:?xt=urn:btih:7742428c76e2bfd26d807cd033bc86d21e5c27a0&dn=Microsoft+Windows+8+Activation&tr=udp%3A%2F%2Ftracker.openbittorrent.com%3A80&tr=udp%3A%2F%2**

Ftracker.publicbt.com%3A80&tr=udp%3A%2F%2Ftracker.istole.it%3A6969&tr=udp%3A%2F%2Ftracker.ccc.de%3A80.



Figura 50. Resultado consulta enlace magnético desconocido.

En este caso, aunque se ha obtenido el enlace del portal The Pirate Bay, no se puede mostrar al usuario información acerca del archivo. Esto puede suceder porque no existe referencia alguna dentro de las bases de datos consultadas o bien porque no se ha rastreado ningún archivo que coincida con la información hash que contiene el enlace magnético.



[Home](#) | [MYPROBE](#) | [Torrent Guard](#) | [Related](#) | [People](#) | [Contact](#)

Check your torrent results

PERFECT! TRUSTED CONTENT

- **Title:** Chelsea.Lately.2012.10.30.Kerry.Washington.HDTV.XviD-AFG
- **ID:** 7781759
- **Date:** 2012-11-01 16:43:26
- **Category:** Video
- **Subcategory:** TV shows
- **Author:** [TvTeam](#)
- **Hash info:** 4ba1a37bd7bea42fd00a41901c4587bfa376fb22
- **Hash info encode:** 4ba1a37bd7bea42fd00a41901c4587bfa376fb22
- **Hash info byte:** K%A1%A3%7B%D7%BE%A4%2F%D0%0AA%90%1CE%87%BF%A3v%FB%22

• **Seeders Information:** the main IPs for this user have been located.

TOP5	IP	ISP	Country	City
1	Anonymized	SK Broadband Co Ltd	We don't know	We don't know
2	Anonymized	Road Runner	US	Westminster
3	Anonymized	Bredbandsbolaget AB	SE	Landskrona
4	Anonymized	University of Tromso, Norway	NO	Troms
5	Anonymized	Com Hem AB	SE	Stockholm
Total IPs: 1185				

• **More Information**

- [User homepage](#)
- [Torrent homepage](#)

[New search](#)

[Home](#) | [MYPROBE](#) | [Torrent Guard](#) | [Related](#) | [People](#) | [Contact](#)

Figura 51. Resultado consulta enlace magnético correcto.

En la figura anterior se muestra el resultado que aparece a través del portal web para un enlace magnético que hace referencia a un archivo correcto (no malware). El enlace es: **magnet:?xt=urn:btih:4ba1a37bd7bea42fd00a41901c4587bfa376fb22&dn=Chelsea.Lately.2012.10.30.Kerry.Washington.HDTV.XviD-AFG&tr=udp%3A%2F%2Ftracker.openbittorrent.com%3A80&tr=udp%3A%2F%2Ftracker.publicbt.com%3A80&tr=udp%3A%2F%2Ftracker.istole.it%3A6969&tr=udp%3A%2F%2Ftracker.ccc.de%3A80.**

La información hash del enlace anterior sí estaba almacenada dentro de la base de datos que consulta Torrent Guard. De esta forma, se puede mostrar información relevante y detallada sobre el archivo en concreto. Además se muestra información de interés sobre el usuario que ha publicado el archivo. Se detallan, como sucede en Publisher Information, las principales direcciones a través de las cuales el publicador realiza su actividad.

También se puede dar el caso de encontrar un enlace cuyo resultado sea positivo, es decir, el archivo al que hace referencia sea correcto. Pero, por el contrario, no se pueda obtener información sobre el usuario que ha publicado dicho archivo. En ese caso, la página indica que no posee información sobre dicho usuario y por tanto no se resumen las principales direcciones que utiliza para publicar torrents: **magnet:?xt=urn:btih:7ce95507b57ad7fecbfe82946af11a6f9f9428ae&dn=Battlefield+3+rus+sound+28.07.2012&tr=udp%3A%2F%2Ftracker.openbittorrent.com%3A80&tr=udp%3A%2F%2Ftracker.publicbt.com%3A80&tr=udp%3A%2F%2Ftracker.istole.it%3A6969&tr=udp%3A%2F%2Ftracker.ccc.de%3A80**. El resultado de esta consulta puede verse en la Figura 52.

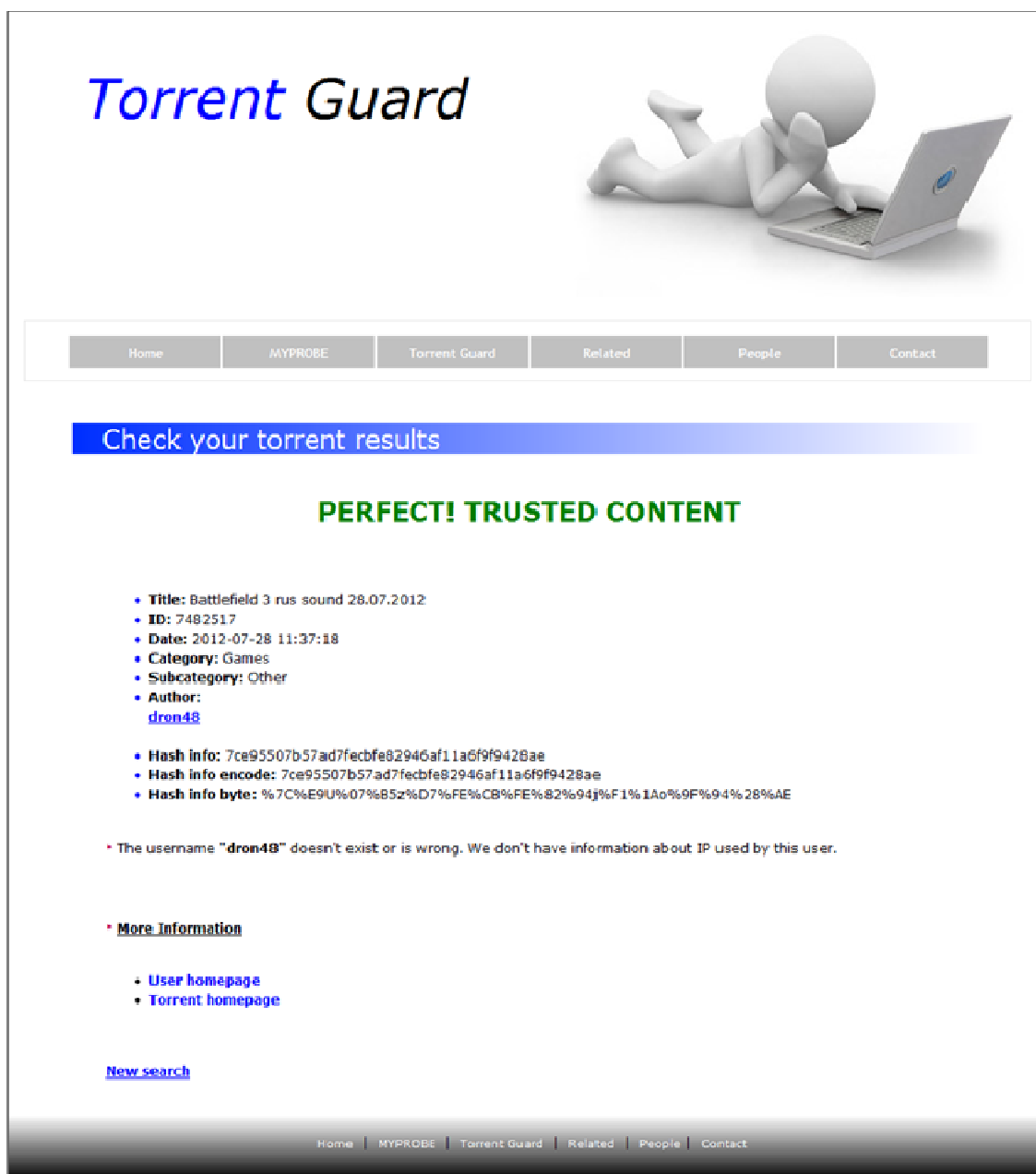
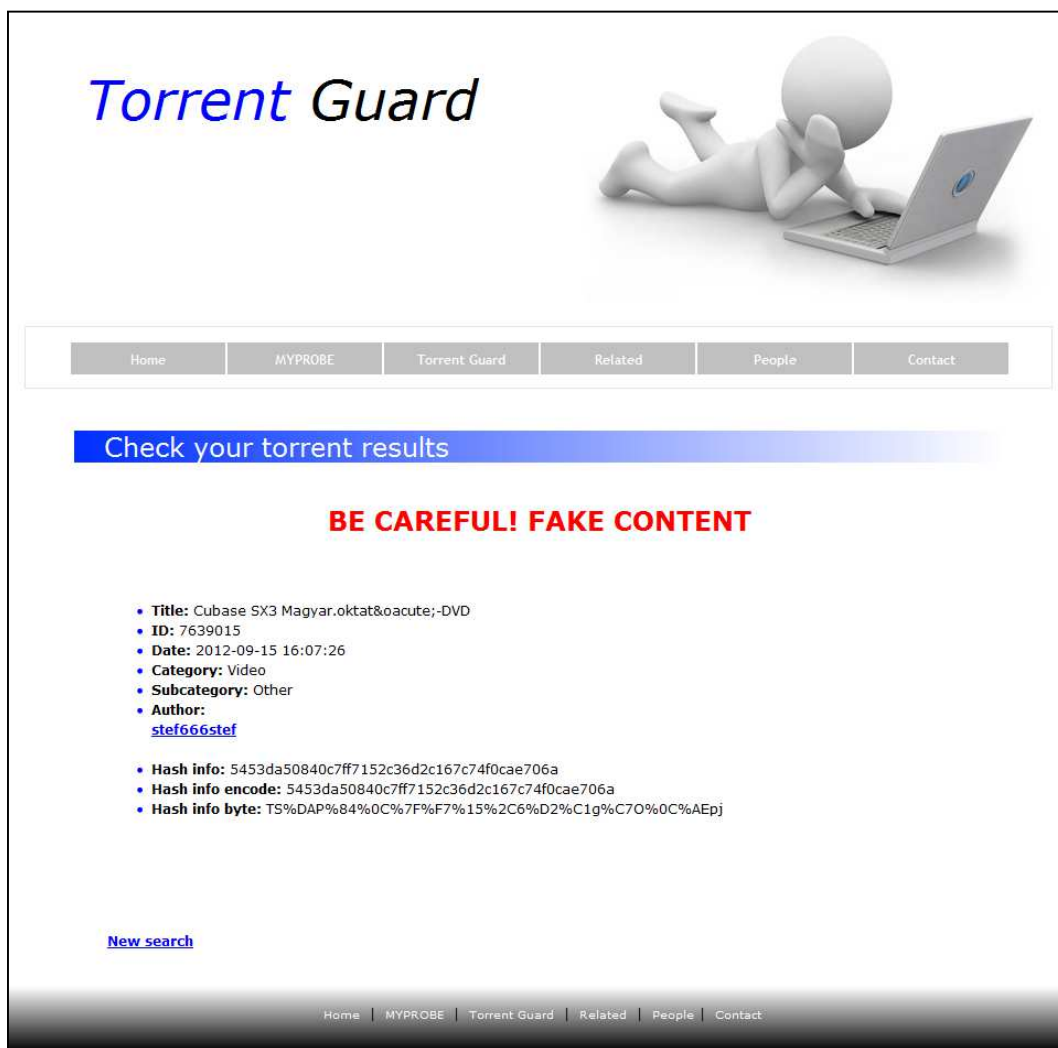


Figura 52. Resultado consulta enlace magnético correcto, sin información de usuario.

Si se consulta un enlace magnético malicioso, es decir, identificado como malware el resultado que aparece en pantalla es el mismo que el de la Figura 43 o Figura 47.

El enlace magnético correspondiente al archivo malicioso analizado anteriormente es: **magnet:?xt=urn:btih:5453da50840c7ff7152c36d2c167c74f0cae706a&dn=Cubase+SX3+Magyar.oktat%26oacute%3B-DVD&tr=udp%3A%2F%2Ftracker.openbittorrent.com%3A80&tr=udp%3A%2F%2Ftracker.publicbt.com%3A80&tr=udp%3A%2F%2Ftracker.istole.it%3A6969&tr=udp%3A%2F%2Ftracker.cc.c.de%3A80.**



The screenshot shows the Torrent Guard website interface. At the top, the logo "Torrent Guard" is displayed in blue. To the right is a 3D illustration of a person lying on their back, using a laptop. Below the header is a navigation bar with links: Home, MYPROBE, Torrent Guard, Related, People, and Contact. A blue banner reads "Check your torrent results". Below this, a red warning message states "BE CAREFUL! FAKE CONTENT". The search results for a magnet link are listed:

- **Title:** Cubase SX3 Magyar.oktató-DVD
- **ID:** 7639015
- **Date:** 2012-09-15 16:07:26
- **Category:** Video
- **Subcategory:** Other
- **Author:** [stef666stef](#)
- **Hash info:** 5453da50840c7ff7152c36d2c167c74f0cae706a
- **Hash info encode:** 5453da50840c7ff7152c36d2c167c74f0cae706a
- **Hash info byte:** TS%DAP%84%0C%7F%F7%15%2C6%D2%C1g%C7O%0C%AEpj

Below the results is a link for "New search". At the bottom, a footer navigation bar contains the same links as the top: Home, MYPROBE, Torrent Guard, Related, People, and Contact.

Figura 53. Resultado consulta enlace magnético malicioso.

CAPITULO 7

COMPORTAMIENTO DEL PORTAL WEB EN DISTINTOS NAVEGADORES

7.1 INTRODUCCIÓN

En este capítulo se describe el comportamiento del portal web en los distintos navegadores con los que es compatible. Inicialmente se realiza un pequeño repaso sobre los principales navegadores utilizados hoy en día.

Al final, se confeccionan algunas estadísticas con el objetivo de conocer el porcentaje de fallos y errores que dan las diferentes páginas del portal, en función del navegador utilizado.

7.2 NAVEGADORES EN EL MERCADO

En la actualidad existen diferentes navegadores web que se pueden utilizar para acceder a Internet. Normalmente los navegadores vienen instalados por defecto en los ordenadores o dispositivos portátiles. Muchas veces, en función del sistema operativo instalado, el usuario tiende a usar un navegador u otro.

Como se observa en la siguiente publicación [35], se pueden elaborar distintas listas con los principales navegadores utilizados en función del dispositivo y versión utilizada. Así, los navegadores de escritorio principalmente utilizados son:

Navegador escritorio septiembre 2012	Porcentaje de uso
Internet Explorer	53.63 %
Mozilla Firefox	20.08 %
Google Chrome	18.86 %
Safari	5.26 %
Opera	1.62 %

Tabla 2. Principales navegadores de escritorio. Septiembre 2012.

El dominio de Internet Explorer es indiscutible, concentrando más del 50 % de los usuarios. Prácticamente, el otro 50 % está dividido entre los navegadores Mozilla Firefox y Google Chrome, estando el primero ligeramente por delante.

Esto refuerza la decisión tomada en el Proyecto creando, diseñando y haciendo el portal web compatible con los principales navegadores.

A continuación se describen brevemente algunos de los aspectos y características fundamentales de los navegadores: Internet Explorer, Mozilla Firefox, Google Chrome y Safari.

7.2.1 INTERNET EXPLORER

El navegador Internet Explorer [36] está desarrollado por Microsoft para el sistema operativo Microsoft Windows desde 1995, estando instalado por defecto en este sistema operativo. Algunas de sus principales características [37] son:

- Elevado nivel de seguridad.
- Permite navegar sin guardar ningún dato de la sesión de exploración.
- Ofrece soporte en las páginas web ActiveX y VBScript.
- Compatible con páginas HTA.
- Sirve de soporte para los *applets*⁽¹³⁾ de Java que funcionan mejor en otros navegadores.
- Permite la descarga de actualizaciones para el propio sistema operativo de Windows.
- Permite editar la definición de accesos directos.
- A partir de la versión 8 incorpora nuevas compatibilidades con el estándar CSS.

Por otro lado, tiene algunas desventajas o limitaciones como:

- No muestra la página web hasta que no finalice la descarga de sus elementos.
- Imposibilidad de acceder a contenidos multimedia por exceso de seguridad.

⁽¹³⁾ Componente de una aplicación que se ejecuta en el contexto de otro programa [38].

- Lentitud de procesamiento de páginas.

7.2.2 *MOZILLA FIREFOX*

Mozilla Firefox [39] es un navegador web libre y de código abierto desarrollado para Microsoft Windows, Mac OS X y GNU/Linux coordinado por la Corporación Mozilla y la Fundación Mozilla. En la actualidad es el segundo explorador web más utilizado.

Sus principales características [37] son:

- Software de código abierto, lo que lo convierte en un navegador totalmente configurable.
- Alto nivel de seguridad.
- Efectiva protección contra el *spyware* y otros tipos de malware.
- Bloqueo asegurado contra *pop-up* y otras formas de publicidad comunes en la web. ActiveX no está permitido por considerarse un riesgo.
- Permite crear y utilizar simultáneamente varios perfiles o preferencias en el mismo navegador.

7.2.3 *GOOGLE CHROME*

Ocupa la tercera posición en el listado de navegadores principales utilizados por los usuarios. Siendo el más joven, Google Chrome [40] es un explorador web desarrollado por Google y compilado con base en código abierto como WebKit.

Entre sus principales cualidades [37] destacan:

- Posee las funciones esenciales y básicas por lo que es ideal para personas con poco dominio en la navegación web.
- Rápido. Emplea un motor renderizado que mejora el rendimiento en la carga de las páginas.
- Google ofrece la actualización automática del navegador.
- Permite navegar sin guardar ningún dato de la sesión de exploración.
- En la página de inicio muestra miniaturas de las páginas visitadas.

7.2.4 *SAFARI*

Safari [41] [37] [42] es un navegador que se ha distinguido por su desempeño, velocidad y soporte de los estándares. Aunque no es tan reconocido para usuarios de otros sistemas operativos diferentes a Mac OS, se ha vuelto una opción interesante desde que salió su versión

para Windows. Como se observa en la Tabla 2, ocupa la 4 posición en el listado de navegadores principales.

Algunas de sus características son:

- Uso de WebKit, como motor de generación de código.
- Campo de búsqueda inteligente.
- Navegación por pestañas.
- Marcadores.
- Bloqueo de ventanas emergentes.
- Permite configurar abreviaciones para el teclado.

7.3 VERSIONES DE LOS NAVEGADORES UTILIZADOS

Para las distintas consultas realizadas en el portal web, se ha utilizado un ordenador portátil para pruebas con distintos navegadores instalados. En este apartado se repasan las diferentes versiones utilizadas en función del explorador.

Navegador utilizado	Versión
Internet Explorer	9.0.8112.16421
Mozilla Firefox	3.6.28
Google Chrome	22.0.1229.94
Safari	5.1.7

Tabla 3. Versiones de los navegadores utilizados.

En el caso de Internet Explorer, se ha comprobado que funciona correctamente sin necesidad de usar la opción Vista de Compatibilidad.

Por tanto, para las versiones mencionadas anteriormente se asegura un comportamiento correcto y completo de las distintas páginas web que forman el portal desarrollado.

Otros navegadores donde se ha verificado el correcto comportamiento del portal han sido:

Navegador	Versión
Safari	6.0.1
Internet Explorer	8.0.7601.17514
Mozilla Firefox	3.6.22

Tabla 4. Otros navegadores compatibles.

7.4 RESULTADOS DEL PORTAL WEB EN DISTINTOS NAVEGADORES

Como se ha explicado anteriormente, la mayor dificultad a la hora de desarrollar un portal web reside en hacer que los distintos navegadores interpreten de la misma forma el código. Para ello, además de entender los ficheros *.html*, también deben interpretar el estilo concreto establecido para las páginas.

En este apartado se va a verificar que el resultado del portal web es el mismo en los distintos navegadores y versiones especificados en el apartado anterior. Comprobando que el aspecto y funcionalidad del portal web coinciden independientemente del explorador utilizado.

Teniendo en cuenta que el portal web está compuesto por 15 páginas independientes con extensión *.html* y 8 archivos *.php* que recogen la funcionalidad del portal, y que además hay que comprobar estos archivos en 4 navegadores diferentes, se van a obtener un amplio abanico de escenarios que se tendrán que comparar entre los exploradores.

Por tanto, realizando un recorrido pormenorizado por el portal web, se van a comprobar los siguientes aspectos:

- Las distintas páginas web se cargan correctamente.
- El contenido de las páginas se posiciona perfectamente.
- El redimensionamiento de la página no supone ningún problema. Es decir, el contenido de la página en cuestión se posiciona correctamente y aparecen los elementos oportunos como, por ejemplo: barras de desplazamiento, que permiten navegar sin problemas.
- En función del explorador que se esté utilizando, se carga la hoja de estilo correspondiente que permite homogeneizar el aspecto de las páginas.
- El icono de la barra de navegación aparece.
- Menú principal:
 - Posicionamiento del mismo.
 - Despliegue en función de la opción seleccionada.
- Pie de página:
 - Posicionamiento del mismo.
- El resultado de las distintas consultas se representan correctamente.
- Los distintos enlaces están disponibles y tienen funcionalidad.

Para poder comparar los resultados obtenidos en los distintos navegadores de una forma organizada, se estructurarán los resultados en una tabla. Se irá comparando una misma característica para los distintos navegadores. Si se obtiene el resultado esperado se indicará con el símbolo ✓. En caso contrario el símbolo utilizado será ✗. Si existe la necesidad de comentar algún resultado, se realizarán las puntualizaciones oportunas.

En la siguiente tabla se resumen las características comunes que deben cumplir todas las páginas. Para contrastar los resultados se ha ido cargando una misma página en distintos navegadores y se han comprobado las distintas características indicadas en la tabla siguiente:

Características comunes en todas las páginas	Internet Explorer	Mozilla Firefox	Safari	Google Chrome	Comentario
Carga de la página web.	✓	✓	✓	✓	Menor tiempo de carga conseguido en Safari.
Posicionamiento correcto del contenido.	✓	✓	✓	✓	
Redimensionamiento.	✓	✓	✓	✓	
Apariencia página web: carga de la hoja de estilo adecuada.	✓	✓	✓	✓	
Icono de navegación.	✓	✓	✓	✓	
Menú: posicionamiento, despliegue y apariencia.	✓	✓	✓	✓	
Pie de página: posicionamiento.	✓	✓	✓	✓	
Enlaces: apariencia y funcionalidad.	✓	✓	✓	✓	

Tabla 5. Características comunes en todas las páginas.

En el caso de aquellas páginas que permiten al usuario introducir información para realizar una consulta, hay que comprobar que independientemente del navegador utilizado se muestran los mismos mensajes de ayuda. Para verificar que las páginas proceden de la misma manera, se cargan las distintas páginas involucradas: Publisher Information, TOP Publishers, TOP ISP, TOP Country y Check your torrent. Se posiciona el ratón alrededor de la zona dedicada a indicar los criterios de búsqueda. Además se prueban los botones “Submit” y “Clean text” introduciendo valores permitidos y otros erróneos. Así se prueba la funcionalidad de las páginas y si el comportamiento y respuesta de las mismas son los establecidos en cada escenario.

Características comunes en Publisher Information, TOP Publishers, TOP ISP, TOP Country y Check your torrent	Internet Explorer	Mozilla Firefox	Safari	Google Chrome
Mensajes emergentes en caso de no introducir ningún parámetro de consulta.	✓	✓	✓	✓
Mensaje informativo al usuario.	✓	✓	✓	✓

Tabla 6. Características comunes en las páginas de consulta.

A la hora de mostrar los resultados, éstos deben presentarse con la estructura y apariencia establecida. Es importante verificar que los resultados de las consultas no sufren

modificaciones al cambiar el explorador. Para probar estas características se realiza una misma consulta en los distintos navegadores.

Características comunes en páginas de resultados	Internet Explorer	Mozilla Firefox	Safari	Google Chrome	Comentario
Cohesión en los datos mostrados como resultado de la consulta.	✓	✓	✓	✓	
Paginación de resultados.	✓	✓	✓	✓	No aplica para Publisher Information y Check your torrent.

Tabla 7. Características comunes en las páginas de resultados.

El portal web da la opción al usuario de introducir distintos criterios de búsqueda para realizar consultas y comprobar el estado de los torrents, pero también ofrece la opción de descargar distintos artículos de investigación y plugins para distintos clientes de BitTorrent. Se ha comprobado que efectivamente se accede a la información correcta en función del enlace seleccionado, independientemente del navegador. Así como que la descarga se produce sin incidentes siguiendo la configuración oportuna en cada navegador.

Características comunes en: páginas con opción de descarga	Internet Explorer	Mozilla Firefox	Safari	Google Chrome
Acceso al material seleccionado.	✓	✓	✓	✓
Correcta descarga del archivo.	✓	✓	✓	✓

Tabla 8. Características comunes en páginas con opción de descarga.

Como sucede con otras páginas que forman el portal, Contact es una página que permite al usuario introducir información. Por tanto no sólo hay que comprobar la funcionalidad de la misma, como se describe en el apartado 5.3.6, es importante verificar que este comportamiento se repite en los exploradores acordados.

Características especiales para Contact	Internet Explorer	Mozilla Firefox	Safari	Google Chrome
Mensajes emergentes en caso de no completar los campos obligatorios.	✓	✓	✓	✓
Envío del correo electrónico: mensaje de agradecimiento.	✓	✓	✓	✓
Recepción del correo electrónico en las cuentas establecidas.	✓	✓	✓	✓

Tabla 9. Características especiales en Contact.

Por lo tanto se puede concluir que todas las páginas que forman el portal web pueden ser accedidas desde distintos navegadores y visualizadas sin incidencias, obteniéndose unas estadísticas muy favorables en cuanto al porcentaje de fallos de acceso que se producen.

	Internet Explorer	Mozilla Firefox	Safari	Google Chrome
Número de páginas accedidas	15	15	15	15
Incidencias producidas	0	0	0	0
Porcentaje de fallos producidos	0%	0%	0%	0%

Tabla 10. Porcentaje de fallos producidos en los distintos navegadores.

CAPÍTULO 8

CONCLUSIONES

El portal web desarrollado en el presente Proyecto cubre por completo los objetivos fijados y explicados en el primer capítulo de esta memoria.

Como propósito principal se establecía ofrecer a los usuarios del entorno de BitTorrent, una forma sencilla y práctica que permitiera verificar el estado de los archivos *.torrent* que se pretendían descargar. Este objetivo se ha conseguido gracias a la aplicación Torrent Guard, que ha sido previamente desarrollada en otro proyecto. El acceso a sus bases de datos brinda la oportunidad de clasificar los torrents en diferentes perfiles en función de su estado y dar a conocer esta información a los usuarios a través del portal web de una forma rápida y directa. Una ventaja añadida es que los usuarios pueden chequear torrents independientemente del portal de descarga del que hayan sido obtenidos. En las bases de datos consultadas por el portal web, generadas por Torrent Guard, también se guarda la información hash asociada a cada archivo *.torrent*. De esta manera también se ofrece la opción de comprobar el estado de los torrents a través de la información hash que los caracteriza, evitando así la descarga previa del propio archivo. Debido a las últimas novedades acaecidas dentro del entorno de BitTorrent, los usuarios de este portal web también tienen la oportunidad de conocer el estado de los archivos *.torrent* a partir de los enlaces magnéticos.

A través del apartado MYPROBE, Monitoring identifying and profiling BitTorrent publishers, se consigue otro de los propósitos de este Proyecto: facilitar a los usuarios la consulta de información significativa acerca de los diferentes publicadores de contenidos dentro de The Pirate Bay. El rastreo de la actividad de estos usuarios se ha llevado a cabo gracias a la aplicación Crawler, igualmente desarrollada en otro proyecto.

Por tanto, el portal web desarrollado unifica el acceso a estas herramientas haciendo que de forma pública se pueda conocer información relevante del ecosistema de BitTorrent.

El análisis de los datos actuales muestra que los principales publicadores dentro de The Pirate Bay son: **TvTeam**, **R0thschild** y **scenebalance**, situados en las tres primeras posiciones respectivamente. El primero comparte un número de torrents bastante elevado: 3558, dentro de la categoría de **Video**. Por su parte, **R0thschild** y **scenebalance** han publicado un total de 1974 y 1562 *.torrent* (rastreados hasta la fecha) principalmente en la categoría de **Porn** y **Video** respectivamente.

El grueso de archivos rastreados por la aplicación Crawler han sido publicados en The Pirate Bay utilizando proveedores de servicio de Internet como **Comcast Cable**, **Ovh Systems** o **Road Runner**. En la actualidad el uso de enlaces magnéticos permite la descentralización de los archivos, es decir, no hace falta que el servidor esté disponible de forma permanente.

Por tanto, el portal web no sólo permite a los usuarios consultar el estado de archivos que pretenden descargar en diferentes formatos, previniendo la descarga de contenidos malware y protegiéndolos. Además ofrece la oportunidad de consultar información acerca de los principales publicadores dentro del ecosistema de BitTorrent.

CAPÍTULO 9

PRESUPUESTO

9.1 INTRODUCCIÓN

Este capítulo resume los costes económicos que hay que afrontar para poder llevar a cabo el Proyecto. Para elaborar el presupuesto se ha utilizado la plantilla disponible en Aula Global.

9.2 PRESUPUESTO

9.2.1 *IMPACTO ECONÓMICO*

Comenzando con los costes directos del presupuesto, en primer lugar se analizan los recursos humanos que han sido necesarios para llevar a cabo este Proyecto. Teniendo en cuenta que han participado 3 Ingenieros, dos de ellos con un perfil *Senior*, se han calculado el número total de horas dedicadas por parte de cada uno. Posteriormente, se ha calculado la dedicación de cada uno de ellos. Este valor está medido en hombres-mes, teniendo en cuenta la relación establecida en la plantilla de presupuesto. Finalmente, se ha calculado el coste total asociado a cada uno de ellos.

En segundo lugar se analizan los costes materiales, tanto en *hardware* como en *software*, que han sido necesarios. A lo largo de todo el Proyecto se han utilizado dos ordenadores, uno de ellos ubicado en el laboratorio 4.1.F04 de la Universidad. El otro, un ordenador portátil

personal. En cuanto al software utilizado, simplemente ha sido necesaria la instalación del programa Eclipse así como las distintas aplicaciones del entorno Microsoft Office. El resto de programas utilizados estaban disponibles en los ordenadores, por tanto, el coste de los mismos ha sido nulo.

La tasa aplicada para el cálculo de costes indirectos es del 20%, como se indicaba en la plantilla de presupuesto.

9.2.2 IMPACTO TEMPORAL

La realización del Proyecto se ha llevado a cabo de manera simultánea con otras actividades. Por lo tanto, pese a tener una continuidad temporal desde marzo de 2011, se han producido distintos periodos en los cuales la disponibilidad personal se ha visto limitada. Algunos de los motivos han sido académicos o laborales. Por este motivo, para calcular el impacto temporal que supone el Proyecto, se han tenido en cuenta aquellos periodos de verdadera productividad en los cuales la dedicación ha sido absoluta.

El impacto temporal asociado a este Proyecto es de 13 meses.


UNIVERSIDAD CARLOS III DE MADRID
Escuela Politécnica Superior

PRESUPUESTO DE PROYECTO

1.- Autor:

Mª Almudena González de la Cruz

2.- Departamento:

Ingeniería Telemática

3.- Descripción del Proyecto:

- Título **CREACIÓN Y DISEÑO DE UN PORTAL WEB CON ACCESO A TORRENT GUARD Y CONSULTA DE INFORMACIÓN DE LOS PUBLICADORES DENTRO DEL ENTORNO DE BITTORRENT**

- Duración (meses) **13**Tasa de costes Indirectos: **20%****4.- Presupuesto total del Proyecto (valores en Euros):**

40.128,10 Euros

5.- Desglose presupuestario (costes directos)**PERSONAL**

Apellidos y nombre	N.I.F. (no rellenar - solo a título informativo)	Categoría	Dedicación (hombres mes) ^{a)}	Coste hombre mes	Coste (Euro)	Firma de conformidad
González de la Cruz, Mª Almudena	-	Ingeniero	11,27	2.694,39	30.365,78	
Kryczka, Michal	-	Ingeniero Senior	0,46	4.289,54	1.973,19	
Cuevas Rumin, Rubén	-	Ingeniero Senior	0,17	4.289,54	729,22	
					0,00	
					0,00	
Hombres mes 11,9				Total	33.068,19	

^{a)} 1 Hombre mes = 131,25 horas. Máximo anual de dedicación de 12 hombres mes (1575 horas)
Máximo anual para PDI de la Universidad Carlos III de Madrid de 8,8 hombres mes (1.155 horas)

EQUIPOS

Descripción	Coste (Euro)	% Uso dedicado proyecto	Dedicación (meses)	Periodo de depreciación	Coste imputable ^{d)}
Ordenador principal (sobremesa)	549,00	100	13	60	118,95
Ordenador apoyo (portatil)	549,00	100	13	60	118,95
		100		60	0,00
		100		60	0,00
		100		60	0,00
				Total	237,90

^{d)} Fórmula de cálculo de la Amortización:

$$\frac{A}{B} \times C \times D$$

A = nº de meses desde la fecha de facturación en que el equipo es utilizado**B** = periodo de depreciación (60 meses)**C** = coste del equipo (sin IVA)**D** = % del uso que se dedica al proyecto (habitualmente 100%)**SUBCONTRATACIÓN DE TAREAS**

Descripción	Empresa	Coste imputable
Total		0,00

OTROS COSTES DIRECTOS DEL PROYECTO^{e)}

Descripción	Empresa	Costes imputable
Software		134,00
Total		134,00

^{e)} Este capítulo de gastos incluye todos los gastos no contemplados en los conceptos anteriores, por ejemplo: fungible, viajes y dietas,**6.- Resumen de costes**

Presupuesto Costes Totales	Presupuesto Costes Totales
Personal	33.068
Amortización	238
Subcontratación de tareas	0
Costes de funcionamiento	134
Costes Indirectos	6.688
Total	40.128

GLOSARIO

CAPTCHA	Completely Automated Public to Tell Computers and Humans Apart
CSS	Cascading Style Sheets
DHT	Distributed Hash Tables
GeoIP	Geolocation IP
HTA	HTML Application
HTML	HyperText Markup Language
IP	Internet Protocol
ISP	Internet Service Provider
P2P	Peer to Peer
PHP	PHP HyperText Pre-processor
RSS	Really Simple Syndication
SQL	Structured Query Language
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
XML	Extensible Markup Language

REFERENCIAS

1. Wikipedia, la enciclopedia libre. <http://es.wikipedia.org/wiki/Peer-to-peer>. Accedido en Diciembre 2011.
2. BitTorrent. <http://www.bittorrent.com/intl/es/>. Accedido en Diciembre 2011.
3. Tutorial BitTorrent. <http://www.ayudabittorrent.com/>. Accedido en Diciembre 2011.
4. The Pirate Bay. <http://thepiratebay.se/>. Accedido Abril 2012.
5. Wikipedia, la enciclopedia libre. http://es.wikipedia.org/wiki/The_Pirate_Bay. Accedido Abril 2012.
6. David Rubio Vidal: "Diseño de una plataforma para el estudio de la publicación de contenidos en BitTorrent". Proyecto Fin de Carrera, Escuela Politécnica Superior Universidad Carlos III de Madrid, 2010.
7. Eclipse. <http://www.eclipse.org/>. Accedido en Abril 2012.
8. Wikipedia, la enciclopedia libre. [http://es.wikipedia.org/wiki/Eclipse_\(software\)](http://es.wikipedia.org/wiki/Eclipse_(software)). Accedido en Abril 2012.
9. Wikipedia, la enciclopedia libre. <http://es.wikipedia.org/wiki/MySQL>. Accedido en Abril 2012.
10. Manual CSS, hojas de estilo. <http://www.desarrolloweb.com/manuales/manual-css-hojas-de-estilo.html>. Accedido en Octubre 2011.
11. Softonic. <http://www.softonic.com/windows/bittorrent>. Accedido en Abril 2012.
12. Wikipedia, la enciclopedia libre. <http://es.wikipedia.org/wiki/HTML>. Accedido en Abril 2012.
13. AulaClic. <http://www.aulaclic.es/html/index.htm>. Accedido en Octubre 2011.
14. Manual de HTML. <http://www.desarrolloweb.com/manuales/21/>. Accedido en Octubre 2011.
15. Guía breve de CSS. <http://www.w3c.es/Divulgacion/GuiasBreves/HojasEstilo>. Accedido en Octubre 2011.

16. WebEstilo. <http://www.webestilo.com/css/>. Accedido en Octubre 2011.
17. Wikipedia, la enciclopedia libre. <http://es.wikipedia.org/wiki/PHP>. Accedido en Abril 2012.
18. Programación en PHP. <http://www.desarrolloweb.com/manuales/12/>. Accedido en Octubre 2011.
19. PHP: Hypertext Preprocessor. <http://www.php.net/>. Accedido en Septiembre 2011.
20. Michal Kryczka, Rubén Cuevas, Roberto González, Ángel Cuevas y Arturo Azcorra: "Torrent Guard: stopping scam and malware distribution in the BitTorrent ecosystem". Artículo de investigación. Instituto IMEDIA Networks, Universidad Carlos III y Telecom SudParis, 2011.
21. Xataka On. <http://www.xatakaon.com/tic/bittorrent-ya-es-el-servicio-protocolo-que-genera-mas-trafico-de-datos-en-europa>. Accedido en Abril 2012.
22. Rubén Cuevas, Michal Kryczka, Ángel Cuevas, Carmen Guerrero y Arturo Azcorra: "Measuring BitTorrent Ecosystem: Techniques, Tips and Tricks". Artículo de investigación. Universidad Carlos III e Instituto IMEDIA Networks, 2011.
23. Internet para todos. http://www.internetparatodos.es/index.php?option=com_content&view=article&id=12560:la-escasa-oferta-legal-impulsa-el-trafico-bittorrent-en-europa-segun-sandvine&catid=36:general-blogs-portales&Itemid=56. Accedido en Abril 2012.
24. Rubén Cuevas, Michal Kryczka, Ángel Cuevas, Sebastian Kaune, Carmen Guerrero y Reza Rejaie: "Is Content Publishing in BitTorrent Altruistic or Profit-Driven?". Artículo de investigación. Universidad Carlos III, KOM Lab TU Darmstadt y Universidad de Oregon, 2010.
25. Hosting y Soluciones de Internet, OVH. <http://www.ovh.es/>. Accedido en Mayo 2012.
26. Wikipedia, la enciclopedia libre. <http://es.wikipedia.org/wiki/OVH>. Accedido en Mayo 2012.
27. Beatriz Calvo González: "Diseño e implementación de un plugin para un cliente BitTorrent mejorado con un sistema de recomendación y detección de contenidos falsos". Proyecto Fin de Carrera, Escuela Politécnica Superior Universidad Carlos III de Madrid, 2011.
28. Wikipedia, la enciclopedia libre. http://es.wikipedia.org/wiki/Tabla_de_Hash_Distribuido. Accedido en Abril 2012.
29. Wikipedia, la enciclopedia libre. <http://es.wikipedia.org/wiki/Magnet>. Accedido en Junio 2012.

-
30. Wikipedia, la enciclopedia libre. http://es.wikipedia.org/wiki/URI_scheme. Accedido en Junio 2012.
31. Wikipedia, la enciclopedia libre. <http://es.wikipedia.org/wiki/URN>. Accedido en Junio 2012.
32. Software y web. Genbeta. <http://www.genbeta.com/a-fondo/las-claves-para-entender-los-enlaces-magnet-el-futuro-del-intercambio-de-archivos>. Accedido en Junio 2012.
33. Wikipedia, la enciclopedia libre. <http://es.wikipedia.org/wiki/RSS>. Accedido en Abril 2012.
34. Wikipedia, la enciclopedia libre. http://es.wikipedia.org/wiki/Proveedor_de_servicios_de_Internet. Accedido en Abril 2012.
35. Desarrollo web. <http://www.desarrolloweb.com/actualidad/ranking-navegadores-septiembre-2012-7518.html>. Accedido en Octubre de 2012.
36. Wikipedia, la enciclopedia libre. http://es.wikipedia.org/wiki/Internet_Explorer. Accedido en Octubre 2012.
37. Navegadores y exploradores web, características y comparaciones. <http://norfipc.com/internet/navegadores-web.html>. Accedido en Octubre 2012.
38. Wikipedia, la enciclopedia libre. <http://es.wikipedia.org/wiki/Applet>. Accedido en Octubre 2012.
39. Wikipedia, la enciclopedia libre. http://es.wikipedia.org/wiki/Mozilla_Firefox. Accedido en Octubre 2012.
40. Wikipedia, la enciclopedia libre. http://es.wikipedia.org/wiki/Google_Chrome. Accedido en Octubre 2012.
41. Wikipedia, la enciclopedia libre. [http://es.wikipedia.org/wiki/Safari_\(navegador\)](http://es.wikipedia.org/wiki/Safari_(navegador)). Accedido en Octubre 2012.
42. Apple, Safari. <http://www.apple.com/es/safari/features.html>. Accedido en Octubre 2012.
43. Comcast Official Site. <http://www.comcast.com>. Accedido en Octubre 2012.
44. Road Runner Domains. <http://www.roadrunnerdomains.com/>. Accedido en Octubre 2012.